

1

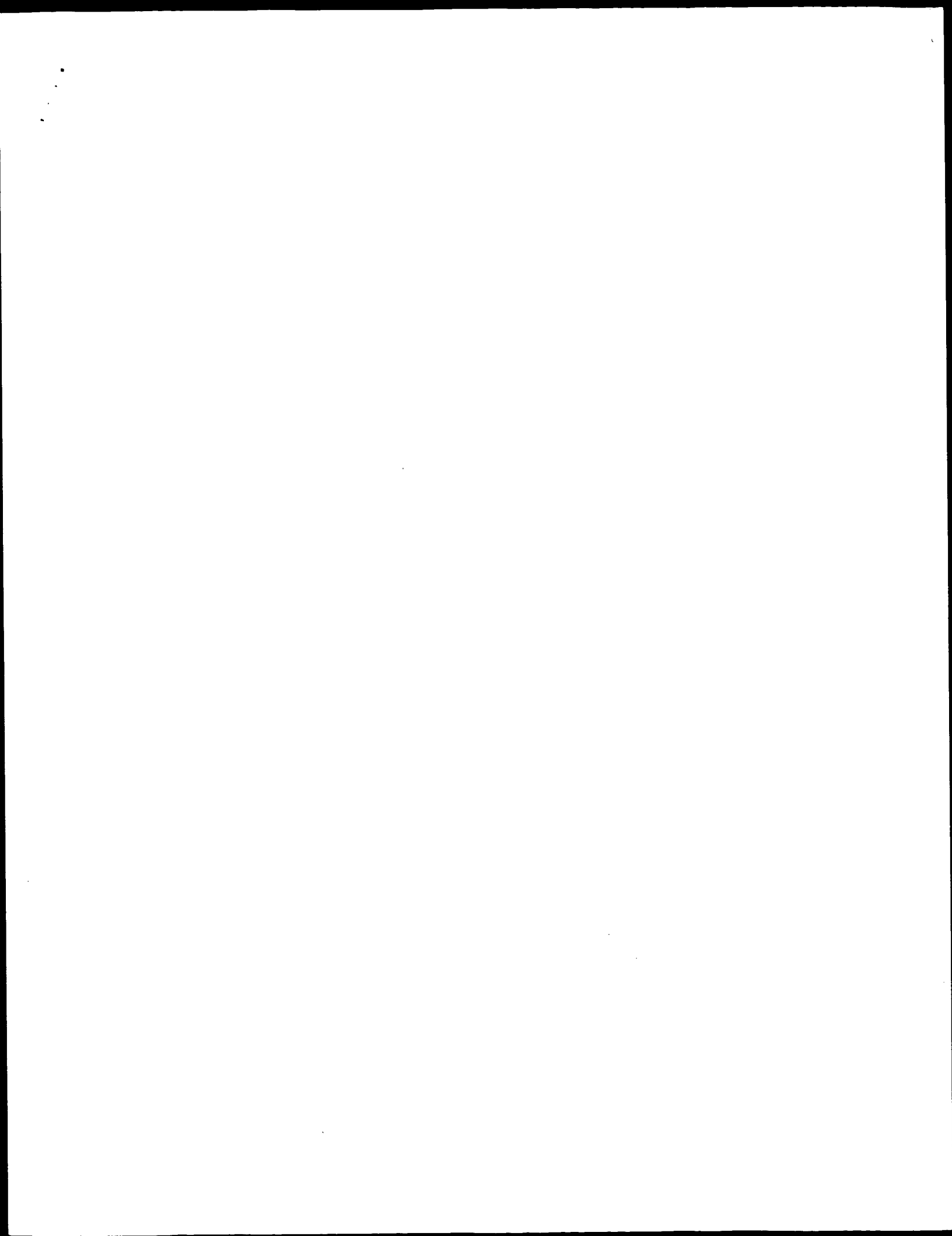
RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463

```

OY      586   gacttgctcaagaaatgttaccgagtacccctgatgcgttcgtcggagaagct    645
          |||
Db      1448  GAACAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR    1389
OY      646   atcgagaatcgtccaagacagattctcgttgagaagccctgcgaacaagatgg    705
          : : : : :
Db      1388  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR    1322
OY      706   ggacagttattccaataaacacctcccgacgggaatgctctcccaccttltat    765
          : : : : :
Db      1338  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR    1265
OY      766   agagacatttccctaataatatagaacatcgtgacacatgaatcattaggagatg    825
          : : : : :
Db      1268  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR    1205
OY      826   ggaagacatcgtttaacagagaattcacctgcgagaagtgaacaagaagaagttactgt    885
          : : : : :
Db      1208  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR    1149
OY      886   ttacagtatcatcagaataatagtaagcgccttcgagacacaacaatcaagatctgg    945
          : : : : :

```

Job time: 13474 sec



/note="ADNC codant pour la proteine humaine BTICP"
70.. 1779
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07740.1"
/db_xref="GI:10043428"
/translation="MDPAVAVLOERALKFPMNSEREDCNNGEPKRIIPKNSLRQY
NSCARCLNOETVCLASTMTKENCVAKIKLANGTSSMTVPKORLTSYKERELECY
KYEOWSESDOYEVEHLISOMCHONGHINSYKIPMLORDPITLTPARGIDHAIEN
LSYLAKSICAEIYCKEWRVTSQGMIMKRIEMVRDLSIMRLAERRGQYLK
NKPDPGNAPNSFTRALYKIIIDJLETSNRCGRHSIQRJHCHSETSKGYCLQYD
DORIVSGLDNNTIKIMDKNTECKRIIIGHTSVLCIODEVRIITGSSDSIVRWYD
NFGEMLNTLIHCEAVLILRFNNGMMVCSKDRSLAVMDASPTDITLRLVJGHRAA
VNVVEPDKYIVASGSDRTIKVMTSTCEFRVLNGHRIACLOYPRLVYSSSDN
TIRLMDIECGACILRVIEGHEELRCIRFENPKRIVSAGDKIKVMDIYAALDPRAPAG
TLRLRTLVHSGRFRVRLQDEFQIVSSSHDDIILIMPLINDPAAQAEPRSRISRYTY
ISR"

BASE COUNT 628 a 467 c 513 g 543 t
ORIGIN

Query Match 100.0% Score 2151: DB 6: Length 2151:
Best Local Similarity 100.0% Pred. No. 0:
Matches 2151: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 tgcgttgctgcgcctgcgcacaaagggcgcccgcgagagcgagaccagtgcc 60
Db 1 TCGGTTGGCTGCGCGCTGGCCCAAAAGGGCGCGCCCGGCGAGACGCCAGTGGCG 60

QY 61 tgcggattatlgagcccgcgagcgcggtgcgcaagaagaagcactcaattgat 120
Db 61 TCGGCGATTATGAGCCCGCGGCGGCGGTCTCAAGAGAGGCACTCAAGTTATGAAAT 120

QY 121 tctcagaagaagaagactgtaataatggaacccccctagagaagataatcagaagaag 180
Db 121 TCCTCAGAGAGAGAAAGACTTATATATGCGAAGACCCCTAGAGAAATATACCAAGAAAG 180

QY 181 aattcattagacagaatacaacagcgtgcagacactcgtcttaaccagaagaacagta 240
Db 181 AATTTCATTAGACAGACATCAACACACTGTGCAGACTGTGCTTAAACCAAGAAACAGTA 240

QY 241 tgtttcagaagacactgtaagaagactggaattgtgtgccaacaaacacttcgcaat 300
Db 241 TGTTTAGCAAGACACTGTATGAAAGACTGAGAAATGTGTGGCCAAACAAACACTTGCCTCAAT 300

QY 301 ggcacttccagta tga tttgtcccaagcaacgaaactcagcaagcattgaaagaagaa 360
Db 301 GGCACCTCCAGTATGATTTGTGCCCAAGCAACGAAACTCTCAGCAGCATATGAAAGAA 360

QY 361 aaggaacgtgtgtcaataactttgacagtggtcagagtcagatcaagtgaattgtg 420
Db 361 AAGGAACGTGTGTCAAAATCTTTGAGCAGTGTGCAGACTCAGATCAATGCAATTTTGTG 420

QY 421 gaacactatatacccaaatgtgtcatcttaacacatgagccatataaccgfatcttaaa 480
Db 421 GAACACTTATATCCCAAAATGTGTCTATTACCAACATGGGCACTAATACTGATCTTAAA 480

QY 481 cctaigtgtcagaagaatttcaactgtctgcagcgtcgaggatltgatatatacgct 540
Db 481 CCTATGTCGAGAGAGATTTCATAACTGCTGTGCCAGCTCGGGGATTTGATATATCGCT 540

QY 541 gagaacaatctgtcatactcgtgatccaaataactatgtgtcgtcgaactgtgtgcaag 600
Db 541 GAACAACATTTCTGATACCTGATGCGCAAAATCATATGTGTGTCGAAGTTTGTGCAAG 600

QY 601 gaatgtaccgaagtagccctgatgagcatgctgtgaaagaactatagagaagatgtgtc 660
Db 601 GAATGTATCCGAGTAGCTCTGTATGCGCATGCTGTGAAAGAACTTATGACAGAAATGTGTC 660

QY 661 aggaagacattctgtggaagagcctgcgcaagaagaagaatgggaagacattatc 720
Db 661 AGGACAGATTCTGTGTGAGAGAGGCGCTGCGCAAGCAAGAGATGGGAGACATATTATTTC 720

QY 721 aaaaaaacacccctctgacgggaatgctctctcccaactcttttatagagacattatcct 780
Db 721 AAAAAAACACCTCTCTGACGGGAATGCTCTCCCACTCTTTTATATAGACACTTATCTCT 780

QY 781 aaaaattatacagaacattgagaacaatagaaatctaatttggagatgtggaagacatagttta 840
Db 781 AAAAAATTATACAAGACATTTGAGACATATAGAAATCTAATTGGAGATGTGGAAGACATAGTTTA 840

QY 841 cagagaatctcctgcgcgaagtgtgaaacaaaggaaggatttactgtttacagtatgatgtat 900
Db 841 CAGAGAAATTCCTGCGCAAGTGAAGCAAGCAAGAGAGTTACTGTTTACAGTATGATGAT 900

QY 901 cagaaaaatgaagaagcgcttcgcagacaacaactcaagatcctggaataaaacacactg 960
Db 901 CAGAAAAATGATGAAGGGCGCTTCGAGACACACATCAAGANTCTGGATTAACAAACACTTG 960

QY 961 gaatgcaagcaattctcaagaccatacaggttcagttcctctgtctccagtatgatgat 1020
Db 961 GAATGCAAGCAAGTTCCTCACAGGCCATACAGTTCACTCTGCTCCAGTATGATGAG 1020

QY 1021 agaatgatacataaagaagatcagatccagatccaggttgagagtgtaataacaggt 1080
Db 1021 AGAATGATCATACAGAGATCATCGGATTCACGGTCAAGAGTGTGGATGTAATATACAGGT 1080

QY 1081 gaaatgtcaaacacglttgatccacatgtgaaagcagttctgcactgtcgtttcaataat 1140
Db 1081 GAAATGTCAAAACAGTTGATTCACCATTTGTAACAGATGTGCACTTGCGTTCAATAT 1140

QY 1141 ggcattgatgtgacctgtctccaagaatcgtccatctgtatagagatagacctccca 1200
Db 1141 GGCATGATGATGAGCTGCTCCAAAAGATCGTTCCATTTCTGTATGAGATGGCTCCCA 1200

QY 1201 actgacattaccctcccggaagggtgtgtgagacacagcgtgtgtcgaatgtttgtagac 1260
Db 1201 ACTGACATTACCCTCCGAGGGTGTGTGTGTGACACACGAGCTGTGTCAATGTTGTATGAC 1260

QY 1261 ttgatgcaagtaacatgtttctgtcatcttgaggatagaactaaagiatggaacaca 1320
Db 1261 TTTGATGCAAGTACATTTGTTGTGCAATCTGGGATGGAACATTAAGATGATGAAACACA 1320

QY 1321 agtactgttgaaattgttaagaaaccttaatggaacaaagaagcattgtctgttgcag 1380
Db 1321 AGTACTGTGAATTTGTATGAGACCTTTAAATGACAAACAGAGCATTTGCTGTGTCAG 1380

QY 1381 tacagggaacagcgtgtgtagttagtgcacatctacacacatcaaatataggaacata 1440
Db 1381 TACAGGGACAGCGTGTAGTGTAGTGCCTCATCTGACAAACATACATATATGGAACATA 1440

QY 1441 gaatgtggtgcatgttttaagagtglttagaaggccaatgaagaattgtgtgtatcgta 1500
Db 1441 GAATGTGTCATGTCATTACAGAGTGTAGAAAGGCCATAGGAATGTGCGCTTGTATTCGA 1500

QY 1501 ttgataaagaagaatgaatgaatgagtgaggccttaatgagaaatlaaagtgaggatctt 1560
Db 1501 TTTGATTAACAAGAGATGATGATAGTGGGCGCTATGATGAAATTAAGATGTGGAGATCTT 1560

QY 1561 gtgcgtgctttgagccccctgtcctcctgcagagggaacactcgtctacagacccctgtggaag 1620
Db 1561 GTGCCTGCTTTGAGACCCCGTGTCTGCTGAGGAGCACTGTGTACGAGACCTTTGGAGAG 1620

QY 1621 catcccggaagagttttctgactaagtttgatgaattccagattgtcagtagttcaat 1680
Db 1621 CATTCGCGAAGAGTTTTCACATCACTTGTATGATGAATTCAGATTTGTCAGTATTCACAT 1680

QY 1681 gatgacaacatccctcatctgagacttcttaaatgaatccagcttcccaagtgtaaccccc 1740
Db 1681 GATGACAACATCCCTCATCTGAGACTTCTTAATGATCCAGCTCCCAAGTGAACCCCC 1740

QY 1741 cgttccctctctcgaacatacacctacatctcagataataaataacatacctaacctcat 1800
Db 1741 CGTTCCCTCTCTCGAATATACACTACATCTCCAGATAAATTAACATATACACTACCTCAT 1800

QY 1801 acttgcgcagagaccataaatttcggtatttaacgtatctgcacaataaccagatgagc 1860
Db 1801 ACTTGCCAGGAGACCCATTAAAGTTGGGGATTAACTGACATTCGCCAATACCGAGATGAGC 1860
QY 1861 aacaacagtaaacatcaactactgcccagttcccttgactagccgagagagcagatgctt 1920
Db 1861 AACAAAGTAACATCAATCACTACTGCCAGTTTCCCTGGACTAGCCGAGAGCAGAGGCTT 1920
QY 1921 tgaagactctgttggagacaglttgctcagctgcgcccagagagcgtctactacagaca 1980
Db 1921 TGAGACTCCTGTTGGGACACAGTTGATGCTGCAGTGGCCGACAGAGGCTACTACGACA 1980
QY 1981 actgactgttgaagtgctgtctatcagaagatgtctctatcaatgtgaaatgttgac 2040
Db 1981 ACTGACTGCTTCACTGCTGTCTATCAGAAATGTCTTATCAATGTGAATGATGGAAC 2040
QY 2041 tttaaacctccctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2100
Db 2041 TTTTAACTCCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2100
QY 2101 cagacaaagtgactataataatatttagtgttttgcagaaaaa 2151
Db 2101 CAGACAAAGTGACTTATTAATATATTAGTGTTCGCCAGAAAAA 2151

RESULT 2

LOCUS AX057166 2151 bp DNA Linear PAT 17-JAN-2001
DEFINITION Sequence 42 from Patent W0075184.
ACCESSION AX057166
VERSION AX057166.1 GI:12309979
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 2151)
AUTHORS Zhang, H., Tsvetkov, L.M. and Kondo, T.
TITLE Modulation of protein levels using the scf complex
JOURNAL Patent: WO 0075184-A 42 14-DEC-2000;
YALE UNIVERSITY (US)

FEATURES
source
1. 2151
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 628 a 467 c 513 g 543 t
ORIGIN

Query Match 100.0%: Score 2151; DB 6; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgcctgcgcacaaaggcgcccgccgagagcgagaccagtgagc 60
Db 1 TCGGTTGGCTGGCGCTGGCACCAAGGGGGCGCCCGCGAGAGCGAGCCAGTGGCC 60
QY 61 tcggagattatgagcccgccgagcggtgtgtcagaagagagcactaaattatgat 120
Db 61 TCGGGATTATGAGCCCGCGAGCGGTGTGTGCAGAGAGGCGCACTCAAGTTATGAT 120
QY 121 tccctcagagagagagactgtataatagcgaaaccccttagaagataataccagagag 180
Db 121 TCCTCAGAGAGAGAGACTGTATATATGCGAACCCTTAGAAGATATATCCAGAGAG 180
QY 181 aattcaacttagacagacatcacacagctgtgccagactctgtcttaaaccaagaaagla 240
Db 181 AATTCACTTAGACACATCACACAGCTGTGCCAGACTCTGCTTAACCAAGAAACAGTA 240
QY 241 tgtttagcaagcaactgtctatgaagaatgtgtgtgccaacaaacaaactgtccaat 300
Db 241 TGTTTAGCAAGCACTGCTATGAGAGCTGAGAAATGTGTGTGSCCAAAACAACTTGCCAAT 300

QY 301 ggcactccagatgaltgttgcgccaagcaacggaactctcagcaagctatgaagaagaa 360
Db 301 GGCACTTCCAGTATGATTTGTGCCAAGCAACGGAACCTCTCAGCAAGCTTGAAGAGAA 360
QY 361 aaggaactgtgtgtccaataactttgaagcagtggtcgaagtcagatcaagatgtgaattgtg 420
Db 361 AAGGAACGTGTGTCAATAATCTTGGACAGTGTCAAGTCAGATCAAGTGAATTTGTG 420
QY 421 gaacatctatcccaaalgtgtctatccaaacaaagggagacataaactctgtatctaa 480
Db 421 GAACATCTTATTCCTCCCAATGTGTCTATACCAACATGGGCATTAATCTGTATTA 480
QY 481 cctatgttcagagagatltcctaactgtctctgcagctcggagatgtgatalatgc 540
Db 481 CCTATGTTGACAGAGATTTCATTAAGTGTCTGTCCAGCTCGGGATGTGATCATGCT 540
QY 541 gagaactctcgtatccctcggatgcgaactcactatgtctgtcgtgaactgtgtgag 600
Db 541 GAGAACATTCGTATACCTGGATGCCAAATCACTATGTCTGTGAACCTGTGTGCAAG 600
QY 601 gaatgtacagagagactctgatgcatgtgtggaagaagctatcagagagatgtgc 660
Db 601 GAATGTACAGAGAGACTCTGATGCGATGCTGTGGAAGAAGCTTATCGAGAGATGTGC 660
QY 661 aggaacagatctctgtgtgagagagcctgtgcagaaagagagatgtggagatattatc 720
Db 661 AGGACAGATTCCTGTGTGAGAGGCTGTGCAGAACGAAGAGATGGGAGATTTATTC 720
QY 721 aaaaaaacctctcgtcggagaaatgctcctcccaactctttatagagacttctc 780
Db 721 AAAAAACAACCTCTGTGACGGGAATGTCTCTCCCAACTCTTTATAGAGCATTTATCCT 780
QY 781 aaaaatlacaaagacatltgagacatagaaatcctaattgagatgtggaagacatgtta 840
Db 781 AAATTTATACAGACATTTGAGACAAATAGAAATCTAATGTGAGATGTGGAACATAGTTA 840
QY 841 caggaatctcagctcggagatgtgaagaagaagagatcttactgttctacagatgatgat 900
Db 841 CAGGAATTCAGCTGCCAAGATGAAAGCAAGAGAAATTTACGTGTTCAGTATGATGAT 900
QY 901 cagaaatagtaagcggccttcgagacacacaacaatcagaatctcggagataaacaacatg 960
Db 901 CAGAAATAGTAAGCGGCTCTTCGAGACACATCAAGATCTGGGATTAATAACACATTC 960
QY 961 gaatgcagcgaatctcagacagccatcacaggttcagctctgtctcagatgtatgag 1020
Db 961 GAATGCAGCGAATTCACAGAGCCATACAGGCTTCAAGTCTGTCTCCAGTATGATGAG 1020
QY 1021 agagtgatcacaagagatcagatcagatcagatcagatcagatcagatcagatcagat 1080
Db 1021 AGAGTGTATCATACAGATCATCGGATTCACAGGCTCAGAGTGTGAGTGAATACAGGT 1080
QY 1081 gaaatgtcaaacagcgtgtgattccacatgtgaaagcagtttgcacttgcgttcaataat 1140
Db 1081 GAAATGTCAAAACAGTGTGATTCACCATGTGGAAGCAGTGTGCACTTGCATTCAATAT 1140
QY 1141 ggcattgtgtgacgcgtctccaaagatcgttccattgtgtatgtgatatgtgcctccca 1200
Db 1141 GGCATGTGTGTGACGCTGTCCAAAGATCGTTCCATTGTGTATGGGATTCCTCCCA 1200
QY 1201 actgacatlaacccctcggagaggtgtgtgtgtgagacacagcgtctctcaatgttgaagc 1260
Db 1201 ACTGACATTTACCTCCCGAGAGGTGTGTGTGTGACACCGAGCTCTCTCATGTGTGAGC 1260
QY 1261 ttgtatgacaglacatgtttctgtcatctgggagatagaactaataaaggtatagaacaca 1320
Db 1261 TTGTATGACAGACATGATTTCTGTGATCTGGGATAGAACTTAAGGATAGAGACACA 1320
QY 1321 agtactgtgaatgttgaagagaccttaaatgtgacacaaagagagatgtcctgtttgag 1380
Db 1321 AGTACTGTGAATTTGTAAGACCTTAATGAAGACAAACGAAGGACTTCCCTGTTGGCAG 1380
QY 1381 tacagagacagcgt 1440

| | | | | | | | |
|-----------------------|------|---|------|---|---------------|-----------|--------------|
| | | | | /function="regulates the stability of IkappaB alpha- and beta-catenin" | | | |
| | | | | /note="contains one F-box domain and seven WD40 domains; forms an SCF with Skp1 and Cull1" | | | |
| | | | | /codon_start=1 | | | |
| | | | | /product="F-box protein Fbw1A" | | | |
| | | | | /protein_id="AAF04464.1" | | | |
| | | | | /db_xref="GI:6164610" | | | |
| | | | | /translation="MDPAEAVLQEKALFKFNSSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNDETVCLASTAMKTEVCATKLANGTSSMIVPKORLSASYEKEELCYKFEQWSESDVEFEVLEHLSQMHYQGHINSYLPKPMLOPDIITALPARDLDAENILSYDAKSLCAELAEKEMRYVSDGMKRLIERMTDSDLMGLAEKRGQYLERKNSPDGNAPNSFYKALPKIIODETIESNMRCGRHSLOJHCRSETSKRYCLOYLDOKIIVSGLRDNTIKIMDKNTLECKRIILGHNIGSVICLOYDERVITITGSSDSVVMVYNGEMLNLLIHCEAVLILRENNCMYATCSKDRSLAYWDMASPTDITRLRYLVGRRAVNVVDFDKITYVSASGDDTIKWNSTOCEFRILNGHRGIACILOYRDLVSSSDNITRLMDIEGACLRVLESHHELVKRCIRPDNKRIVSAGADGKIWADLVAALDPAAPGTLCLRTLVEHSGRVFRLOFDEFQIVSSSHDDTILIMPLNDPAQAQAEPPRSPRTYVISR" | | | |
| BASE COUNT | | | | 628 a | 467 c | 513 g | 543 t |
| ORIGIN | | | | | | | |
| Query Match | | | | 100.0%; | Score 2151; | DB 9; | Length 2151; |
| Best local Similarity | | | | 100.0%; | Pred. No. 0; | | |
| Matches 2151; | | | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| OY | 1 | tcgcttgctgcgagcctgcgacccaagggcgcccgcgagagagcgagcccgatggcc | 60 | | | | |
| DB | 1 | TCGGTTGGCTGGCTGGCCACCAAGGGGGCGCCCGCGGAGAGCGACCCAGTGGCC | 60 | | | | |
| OY | 61 | tcggcgattatlgaccgcgagcgcggtgctgcgaagaaagcaactcaattatgac | 120 | | | | |
| DB | 61 | TCGGCGATTATGAGACCGCGCGGGGTGCTGCAAGAAAGCACTCAAGTTTATGAAT | 120 | | | | |
| OY | 121 | tcctcagaagaagaagactgtaataatgycgaaccccgtaggaagataataccagaag | 180 | | | | |
| DB | 121 | TCCTCAGAGAGAAAGACTGTATATATGCGCAACCCCTAGAGAAATATVCCAGAGAG | 180 | | | | |
| OY | 181 | aattcaactagaagaacatacacacagctgtagcagactctgtttaaccaagaacagta | 240 | | | | |
| DB | 181 | AATTCACCTTAGACAGACATACAAACAGCTGTGCCACACTCTGTTAAACCAAGAAACGTA | 240 | | | | |
| OY | 241 | tgatttagcaagcaactgctatagaagactgtagtctggtccaaacaaacttggcaat | 300 | | | | |
| DB | 241 | TGTTTAGCAAGCACTGCTATGTAAGACTGTGAATGTGTGCCAAACAAACTTGTCCAA | 300 | | | | |
| OY | 301 | ggcacttcagatgatattgtagcgaagcaagcaactctcagaacagctatagaagaagaa | 360 | | | | |
| DB | 301 | GGCACTTCAGATGATATTGTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGAA | 360 | | | | |
| OY | 361 | aaggaactgtagtcaaatctttagcagtagtgcagatcagatcaagtgaaattg | 420 | | | | |
| DB | 361 | AAGGAACCTGTGTCAAAATACTTTGAGCAGTGTGAGATCAGATCAAGTGGAAATTTGTG | 420 | | | | |
| OY | 421 | gaactcttatatcccgaatgtagtcaataacaaatgggcacataactcgtatcttaa | 480 | | | | |
| DB | 421 | GAACATCTTATATCCCAATGTGTATTAACCAATGGGCACTAACTCTATCTTAAA | 480 | | | | |
| OY | 481 | cctaagttgcaagaagatttcaactgctgtagcagctcgagagtagtgcataatgct | 540 | | | | |
| DB | 481 | CCTATGTTGCAAGAGATTTCATTAATCTGCTGCCAGCTCGGGGATTGATCATATATGCT | 540 | | | | |
| OY | 541 | gagaaactctgtagtcaactgtagtccaaatcaatctgtagtgcgaactgtagcaag | 600 | | | | |
| DB | 541 | GAGAACTTCTGTCTATACCTGGATGCCAAATACATATGTGTGCAACTTGTGTGCAAG | 600 | | | | |
| OY | 601 | gaatgtagtcgagtagcctctgtagtctgtagtgcgaagaagctatccagaagaatg | 660 | | | | |
| DB | 601 | GAATGTTACCAAGTACCTCTGTATGGCATGTGTGGAAGAACTTATCGAAGAAATGCTC | 660 | | | | |
| OY | 661 | aggagaattctctgtagagagagcctgcgaagaagaagatggaacaatattatctc | 720 | | | | |
| DB | 661 | AGGACAGATTCTCTGTGAGAGAGCCTTGCGACAGCAAGAGAGATGGGACATTTATTC | 720 | | | | |
| OY | 721 | aaaaacaacccctcgaacggaatgtagtctctcccaactctttatagaacttactc | 780 | | | | |
| DB | 721 | AAAAACAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTTATCCG | 780 | | | | |
| OY | 781 | aaaattatacagaacatctgagacaataatgaatcttaattggagatgtagaacaat | 840 | | | | |
| DB | 781 | AAAATTTATACAGACATTGAGACATATGAATCTTAATTGGAGATGTGGAACATATAGTTTA | 840 | | | | |
| OY | 841 | cagaagaattcactccgaatgtaacaagaagaagattcactgtagttagatgtagat | 900 | | | | |
| DB | 841 | CAGAGAAATTCACCTCCGAAGTAACAAAGCAAGAAAGATTACTGTTTACAGTATGATGAT | 900 | | | | |
| OY | 901 | cagaanaatagtagcggccttcgagacaacacaatcaagaatcgtggaataaacaacat | 960 | | | | |
| DB | 901 | CAGAAATATAGTAACCGGCTTCGAGACAAACATCAATCAAGATCTGGGATTAACACATTCG | 960 | | | | |
| OY | 961 | gaatcgaacggaattctcagaagccatacagaagttcactgtagtccagatgtagat | 1020 | | | | |
| DB | 961 | GAATGCAAGCAATTTCTACAGAGGCGCATACAGTTCAAGTTCCTCTGTCTCCAGTATGATGAG | 1020 | | | | |
| OY | 1021 | agaatgtagtcaatacagaatcactcgaatccaggtcagagtagtggagtagtaatacaggt | 1080 | | | | |
| DB | 1021 | AGAGTGAATCATTAACAGATCATCGGATTCACAGTCCAGATGTGGATGTAAATACAGGT | 1080 | | | | |
| OY | 1081 | gaaatgctaaacacgtttagatccacattgtgaagcagttctgactctggttcaataat | 1140 | | | | |
| DB | 1081 | GAAATGCTAAACAGTGTGATTCACCATTTGAAGCAGATTTGTCATCTGGTTCAATTAAT | 1140 | | | | |
| OY | 1141 | ggcattgtagtgcagctcctcgaagaatcggttccatctgtagttagatgtagcctccca | 1200 | | | | |
| DB | 1141 | GGCATGATGAGTGCAGCTCTCCAAAGATCGTTCCATTCCTGTATGGGATATGGCTCCCA | 1200 | | | | |
| OY | 1201 | actgaacatccctccggaaggtgctggtgtagaacacgagctgtagtcaatgtagaac | 1260 | | | | |
| DB | 1201 | ACTGACATTAACCTCCGAGAGGTGCTGTGCGACACCGAGCTGCTCAATGTTGTAAGAC | 1260 | | | | |
| OY | 1261 | tttagatgaagaatgacattgtagtctgtagcctgggagtagaactaataagtagaaca | 1320 | | | | |
| DB | 1261 | TTTGAATGACAAAGTACATTTGTTTCTGCATCTGGGATGAGACTATTAAGATATGGAACACA | 1320 | | | | |
| OY | 1321 | agtaactgtagaattgtagaagccttaaatgtagacaacaaagagcaatgctgtagtgcag | 1380 | | | | |
| DB | 1321 | AGTACTGTGAATTTGTAAGAGCACTTAATTAAGACAAACAGAGCATGCTGTTTGCAG | 1380 | | | | |
| OY | 1381 | tacaaggacagagctgtagtagtagtagtctatctgacaacatactcagatataaggacata | 1440 | | | | |
| DB | 1381 | TACAGGACAGAGCTGTGATGAGTGTGCTCATTTGGAACAACATATGAGATTATGGACATA | 1440 | | | | |
| OY | 1441 | gaatgtagtgcagttttagcagaggttagaagggcagtagaagattgtagtgcattatcga | 1500 | | | | |
| DB | 1441 | GAATGTGTGTCAGTTTATACAGTGTATGAAGGCCAATGAGAAATGTGTGTATTCGA | 1500 | | | | |
| OY | 1501 | tttagataacaagaagtagtagtagtggggtcctatgtagtgaataaataaagttaggaatct | 1560 | | | | |
| DB | 1501 | TTTGAATMAAGAGGATAGTCACTGGGGCTTATGATGAGAAAAATTAAGTGGGATCTT | 1560 | | | | |
| OY | 1561 | gttagctgtagttagaccccgtagtctcctcgaaggaacactgtagtagcagcccttggag | 1620 | | | | |
| DB | 1561 | GTTGGCTGTGTGAGCCCGCTGCTCTCGAGGGACACTGTGTATGAGCCCTTGTGGAG | 1620 | | | | |
| OY | 1621 | cattccggaagagtttctcgaactacagtttagaattcagaattgtagtgcagtagtcaat | 1680 | | | | |
| DB | 1621 | CATTCCGGAAGAGTTTGTGAGTACAGTTTGAATGAATTCACATTTGTGATGTTACAT | 1680 | | | | |
| OY | 1681 | gatgacaacaatcctcactcgtggagcttccaaatgtagcgaagtgcgaaccccc | 1740 | | | | |
| DB | 1681 | GATGACAAATCCTCATCTGTGGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCCC | 1740 | | | | |
| OY | 1741 | cgttcccttcggaacatacactacatctcagaataaataacatcacatgacat | 1800 | | | | |
| DB | 1741 | CGTTCCCTTCTCGAACATACACTACATCTCCAGTAAATTAACCATACACTGACCTCAT | 1800 | | | | |

| | | | |
|---|------|---|------|
| D | 271 | GAACAGATGATGTTTAGCAAGACACTGCTATGAGACGTGAGAAATTGTGGGCAAAACAA | 33 |
| O | 292 | cttgcgaatctgcaacttccagtaataatgtatgtgtcccaagcaacgcaactctcaagcat | 351 |
| D | 331 | CTTGCAATGGCACTTCACGATATGTTGTCACCAAGCAACGAAACCTTCAGCAACCTAT | 390 |
| O | 352 | gaaagagaaagagacatgtgtcaaatactttgaagcaatgtgcagatgaatgaatgaatg | 411 |
| D | 391 | GAAGAAGAAAGAAAGCACTGTGTCAAACTTTGAGCAGTGGTCAAGATCAGATCAAGTG | 450 |
| O | 412 | gaattgtggaacalcttatatcccaaatgtgcatatcccaaatgtggacataaactg | 471 |
| D | 451 | GAATTTGGAGACATCTTAAATATCCAAATGTGCATATACCAACATGGGACATAAACCTCG | 510 |
| O | 472 | tatcttaaacctatgtgtgcgaagaagatttaataactgtctctgcgaactcgggatttgat | 531 |
| D | 511 | TATCTTAAACCTATGTTGTCAGAGAGATTTATATACGTCTGTGCACCTGGGGATTTGAT | 570 |
| O | 532 | catatcgtcgtgaacatctctcatactgtgatcgaactgaactactatgtctgtgaact | 591 |
| D | 571 | CATATCGCTGAGACATTTCTGTATACCTGAGTGCCAATCACTATGTCTGCTGATCT | 630 |
| O | 592 | gtgtgcgaagaaatgtgtaccgagtgacctctgtatgtgcattgtgtggaagaagcttaccag | 651 |
| D | 631 | GTGTGCAAGCAATGTCTCCAGATGACCTCTGATGGGATGCTGTGGAAGAAGCTTTATCGAG | 690 |
| O | 652 | agaaatgtgcgaagacagatctctctgtgtgagagcgcttcgcaagaagaagaatgtggacag | 711 |
| D | 691 | AGAAATGGTCAGACAGATTTCTGTGTGAGAGGCTTGCAACAGAGAGATGGGACAG | 750 |
| O | 712 | tattatctcaaaaacaaacctcctgcacggaatgtctctcccaactctttatagaaga | 771 |
| D | 751 | TATTTATTTCAAAAACAAACCTCCTGACGGGAAATGCTCTCCAACTTTTATAGAGA | 810 |
| O | 772 | cttatccctaaatctatcacagaacatttgaagcaatagaatctaattgagatgttgaaga | 831 |
| D | 811 | CTTTATCTCTAAATTTATACAGACATTTGAGACATATGAATCTATATTTGAGATGTGAGAA | 870 |
| O | 832 | catatgttcaagagaattctacgtccggaagtgaacaagaagaagatttactgtttacag | 891 |
| D | 871 | CATATGTTTACAGAAATTTCACTGCGCAAGGAAGAAAGAAAGATTTACTGTTTACAG | 930 |
| O | 892 | tatgatgatcagaanaatgaatgaacgagccttcgagacaacaacaataaagcttcggatataa | 951 |
| D | 931 | TATATATATCAAAATATGATGAGGGGCTTCGAGACACACATCAAGATCTGGATTTAA | 990 |
| O | 952 | aacacattggaatgcgaagcgaaatctctcaacgagcatacagtttcaatctctctccag | 1011 |
| D | 991 | AACACATTTGGAATGCAGGCAATTTTCACAGGCCATATACAGTTCAAGTCTCTGTCCAG | 1050 |
| O | 1012 | tatgatgaagagatgtatcataaacaagatatatcgatcttcaacggtcagagttggagatga | 1071 |
| D | 1051 | TATATATGAGAGAGTACATATACAGGATCATCTCGGATTTCCACGGTCAAGATCTGGGATGA | 1110 |
| O | 1072 | aatacaggtaaabaatgcctcaaacacgcttgatctacacatttgaagcagtttgcgaactgcgt | 1131 |
| D | 1111 | AATACAGGTGAATAGCTAATACACAGCTTGATTCACACATTGTGGAAGCAATTTCTGCATTCGCT | 1170 |
| O | 1132 | tccaataatgtgcatagtatgtgacccgtctccaagaatcgtttccattgtctgatagtgatalg | 1191 |
| D | 1171 | TTTCATATATGGCATGATGAGTGGACCTCTCCAAAGATCTTTCATATGCTGTATGGGATATG | 1230 |
| O | 1192 | ggcccccacaactacatactaccctccggagagtgctgtgtggaacacgagctgtctcat | 1251 |
| D | 1231 | GGCTCCCACTGACATTTACCTCTCGGAGGGTCTGTGTGGACACCGAGACTCTGTGCAT | 1290 |
| O | 1252 | gtgtgtagacttgcatacgaagtacatgttctctgcacacctgggagatagaactataaagta | 1311 |
| D | 1291 | GTGTGTACCTTTATGATGACAGTACATTTGTTTGTGCAATCTGGGATATAGAACTTAAAGTA | 1350 |
| O | 1312 | tggagaacaagatctgttgaatttttgaagaagacttaaatgtgacacaacaacgagatctgc | 1371 |
| D | 1351 | TGGAGACAAAGTACTGTGTGAATTTGTGAGGACCTTAAATGGAACACAAAGAGGACTTGTCC | 1410 |

| | | | |
|----|------|---|------|
| OY | 1804 | ttccacgaccacattaaagtttgcggtatttcaagcatctgcgaataaccagatgttgcaac | 1863 |
| Db | 1741 | ttccccagacacgtttaaagttgcgaattattttaaCAGACCTGCCAACAACGAGATGAAAC--- | 1797 |
| OY | 1864 | aacagtaacaatcaaacactactgccagtttccctcg--gactagccgagagacgagcttt | 1921 |
| Db | 1798 | -----AACATTCAAACACCTCAACCCGGATTTCCCGGACGGAATGAGGAGGACGAGGGCTTT | 1851 |
| OY | 1932 | gagactccctgttggacacagtttgcctgcagtcgcccagagacgagcttaactcacacaa | 1981 |
| Db | 1852 | GAGACTCCTGTTGGGACACAGTCGGTTCAGACGCCGA--CGAGACGGCCTGCTCGGACAC- | 1909 |
| OY | 1982 | ctgactgtcttaagtgctgctatcagaagaatgtctctacaaatgttgatgttgaact | 2041 |
| Db | 1910 | --GGTGGCTCAAGTCTCTCTATCAGAAAGATGCTT--TATCTTGTGTGAATGATTTGCAACT | 1966 |
| OY | 2042 | tttaaaccctccctccctctcc---tcccttcaacctctgcacctagtttttcccaatt-gg | 2097 |
| Db | 1967 | TTCAAGCCTCCCTCCCTCTCCCTTCCCTCCCTCCCTCCCTGACACTGTTTTCCTCCCAATTGGG | 2026 |
| OY | 2058 | ttccgagaaagtgtaactataaataatatttagtgttttgcagagaaaaaataa | 2151 |
| Db | 2027 | TTTCAGACAAAGATCACTTATTAATAATATTATTAGTGTTTACAAAAAATAAAAAA | 2080 |

| AF110396 | LOCUS DEFINITION | 1979 bp | RNA | linear | ROD 28-JUN-1999 |
|----------|--|---------|-----|--------|-----------------|
| AF110396 | Mus musculus beta-transducin repeat-containing protein mRNA, complete cds. | | | | |

| | | |
|-----------|--|------------|
| ACCESSION | AF110396 | |
| VERSION | AF110396.1 | GI:5230821 |
| KEYWORDS | | |
| SOURCE | house mouse. | |
| ORGANISM | Mus musculus | |
| | Euarctopota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1979) | |
| REFERENCE | Winston, J., Elledge, S. J. and Harper, J. W. | |
| AUTHORS | Direct Submission | |
| TITLE | Submitted (01-DEC-1998) Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA | |
| JOURNAL | | |
| ENTRY | | |

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. 1979 |

```

CDS
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    1..1710

```

```

/note="beta TRCP"
/codon_start=1
/product="beta-transducin repeat-containing protein"
/protein_id="A041025.1"
/db_xref="GI:5230822"
/translation="MDPAAVLQEKALKFMNSSEEDCQNNDEPPKIIPEKNSLRQTY
NSCARLINDQETVCLISTPAKTEKNCVAKLANGSSIVTPKORKLSISTEKELQ
KIFQWSSDDVEFEVEHLIISOMCHYOHINSLYKMLQRFITLAPRGDHLITEN
LSYIDAKSLCAELVCKEMWYRTSDGMKMLKILEMWYDITLRLARAKMGQITL
NKPDENAPMNSFYRALPYKLIIDLETIESHWGCRHSLQIHCSSETSKGYCLQY
DOKIVLSGRDNTIKIMDKSTLEQKRIILGHTSSVGLQYDEBRIITGSSDEVRWD
NAGMELPTLTHHCDAVLIHLRFNNGMWTGCSDRSLAVDMASPTDITLRYLVGRR
VNVVDEDFDKYIVSASGDRITKYWNMSTCEFPRTNGHKGAGLQYRLRLVVGSSD
TILRLMDIECGACILVLEHGEHLVCRIFPDNKRIVSGRCDKIKWMDLMAALDPRAP
TLCLRLTVLHSGRVRFLQDFEFQIYSSSHDTIILMDPLNPAHAIEPPRSKITYT
ISR"

```

| | | | | |
|-----------------------|--------------|--|-----------------|--------------|
| Query Match | 72.9% | Score 1567.2; | DB 10; | length 1979; |
| Best Local Similarity | 88.5%; | Pred. No. 0; | | |
| Matches 161; | Conservative | 0; | Mismatches 213; | Indels 16; |
| | | | | Gaps 5. |
| QY | 70 | atgagcccgccgaagcgctgcgaagaagcactcaagtttatgaatccccaag | 129 | |

| | | | |
|----|------|--|------|
| Db | 1 | ATGACCCGGGCACAAAGGGGTGTCGACAGGAGAAAGGCGCTTAATTATGTAATTCCTCAAG | 60 |
| Qy | 130 | agagaagacgtatataatgvgagaccccttagaagaataatccagagaagaattcactt | 189 |
| Db | 61 | AGAAAGACTGTATATATGCGAAGCCCTAGAGMATATATCCAGAGAAATTACCTT | 120 |
| Qy | 190 | agacagacatacaacagctctgcagacatctcttaaaccaagaacagatglttagca | 249 |
| Db | 121 | AGACAGACTTACAAACAGCTGTGCCAGGCTTTGCATTAACCAAGACAGATATGTCTACA | 180 |
| Qy | 250 | agcacctctatgaagactvgaaattgtgtgvgccaaacaaacttggccaatgycactcc | 309 |
| Db | 181 | AGCACTCTTGAAGAGCTGAAATTTGTGTGGCCAAAGCCAAACTTGCATATGCACTTCC | 240 |
| Qy | 310 | agatagattgtgcccaagcaacgaaactctcagcaagctatgaaagaaagaaagactg | 369 |
| Db | 241 | AGCATGATTTGGCCCAAGACGGGAAACTCTACGACAGTATGACAGAGAAAGAGCTG | 300 |
| Qy | 370 | tggtgcaaatctttgagcagtggtgcagagtcacatcaagtggaattgttgaacattc | 429 |
| Db | 301 | TGTTCTCAAGTATTTTGGAGCGTGTGTCCGAGTCTGATCAAGTGGAAATTTGTATACACCTT | 360 |
| Qy | 430 | ataccccaaatgtgtcatctaacacatgvgacataaacctgtatctttaaacctatgtg | 489 |
| Db | 361 | ATAATCCCAATGTCTACTACACAGCATGGGCACATCACTTACCTTAACCTTATGCTG | 420 |
| Qy | 490 | caagagagattcltaactgtcttcgcacgtccgggagttgatacatatcgtctgagacat | 549 |
| Db | 421 | CAGAGGGATTTCATTAAGTGCACCTGCCAGCACGGGGCTGTGACCATCACTGAGAACAT | 480 |
| Qy | 550 | ctgtcataactgtgtagccaaatatactatgtgtcgtcgaactgttgtgaagaagtgtac | 609 |
| Db | 481 | CTGTCAATCTTGGAGCGCAAGTCACTGTGTGCTGTGACCTGTGTGCAAGAAATGTATC | 540 |
| Qy | 610 | cgaatgacctctgtatgtagcatgctgtgtgaagaagcttatcgagaagtgtgtcagacaag | 669 |
| Db | 541 | CGCGTAGCTCCGACGGGCATGTGTGGAAACCTCATGCAGAGGATGTGTGAGCGGAC | 600 |
| Qy | 670 | tctctgtggaagagcctcgtgcagaagaagagagtgvggacagtatattatcgaaaaaaa | 729 |
| Db | 601 | TCTCTGTGGCCAGGCTGTGGAGAGCGCAGGCGTGGGAGACACTTATTCAAAAACAA | 660 |
| Qy | 730 | ccctctgacggaaatgtctctcccaactcttcttatagacactatatacctaanaatta | 789 |
| Db | 661 | CCTCTGATAGGAAGCGTCTCTCCCAATCCTTTATATAGAGCGTTATTCCTAAATCAT | 720 |
| Qy | 790 | caagaacattgagacaatagaatctaaatttvgagatgtggaagacataagttcacagaaat | 849 |
| Db | 721 | CAAGACATTTGAGCAAAATAGATGTCCAATTTGAGATGTGTGGCGACATAGTTTACAGAGATC | 780 |
| Qy | 850 | cacgtgcgaagtggaaacaagaagaagttactgttttcacgatatagtataagaanaata | 909 |
| Db | 781 | CACGTCCCGAGTGAACAAAGTAAAGGGGTTACTGTTTACAGTACAGACACAGAGATTA | 840 |
| Qy | 910 | gtaaagcgaccttgcagacaacaacatcaaatcttggatataaanaacacatttgaatgcaag | 969 |
| Db | 841 | GTCAGCGGCGCTTCAGAGCAACACCATCAAGATCTGGGATATAAAGCACACTGGAAATGCAAG | 900 |
| Qy | 970 | cgaattctacagcacaatacagttcaagtcctctgtctctccagtatgataagaagatgac | 1022 |
| Db | 901 | CGGATTTTCACGGGGCCACACGGGCTCGTCTCTGTCTCCAGTACCATATGAGAGGCTATC | 960 |
| Qy | 1030 | atacagaatataatcgattcttcacaggtacagatgttggatagttaataacagtgaaatgcta | 1088 |
| Db | 961 | ATCACAGGCTCTTAGACTCCACCCTGACAGTGTGGGATGTAAATCCAGTGAATGCTTA | 1022 |
| Qy | 1090 | aacacgttattatccaatgttgaagcagttctgcacttgcgttttcaataatgycatgatg | 1148 |
| Db | 1021 | GACACATGTAGTCTACCACTGTGAAGCGTGTCTCCACTCGGCTTCATTAATGTGCATGTATG | 1086 |
| Qy | 1150 | gtgacgtgctcacaagaatglttccatgtgcatgtgataatggatataggccctcccaactgacat | 1209 |
| Db | 1081 | GTGACCTGTTCAAAAGACGCTTCATCGCTGTGTGGGATATGTGCTTCCCAACTACATC | 1140 |

QY 1210 accctcggaggggtgctgctgcagacccagctgctcgaatgctgtaacttgcac 1269
 DB 1141 ACCCTCAGAGAGGGTGTCTGTGGGACACCGAGCTGGGTCAATGTTGTATGATGAC 1200
 QY 1270 aagtaactgttctcgcactcgtggatagactaaagatagtaacacaaactgtt 1329
 DB 1201 AAGTACATGTTCTGCTCCCTGGAGATGACCAATGATGAGTGTGACACCAAGTACCTGT 1260
 QY 1330 gaattgttaaggaaaccttaattgacacaaagagcattgctgttggatagcaggac 1389
 DB 1261 GAATTGTGTAAGACCTTAATGAGCAACAGCTGGCATGCTGTTGTGACATGAGAGAC 1320
 QY 1390 aaggtgttagttagtgcactcgtacacaaactacatcaatlatagggacataaagtgt 1449
 DB 1321 AGGCTGTGTGTGGGCGCTCCTGTGACAAACACATCAGGCTGTGGACATAGAGTGTGA 1380
 QY 1450 gcatgttctcagagtgcttagaaggcattgaggaattgtgtgttattcgaattgatac 1509
 DB 1381 GCATGCTCCTCGAGTGTGGAGGGCATGAGGATGTAACGCTGCATTGATTTGATTAAC 1440
 QY 1510 aagaagatagctagtgaggccatgatgataaataaagtgtgagctgtgtgtgt 1569
 DB 1441 AAAAGATAGTGTAGGAGGAGGAGCTATGATGGAATAATTAAGTGTGGATCTTATGCTGT 1500
 QY 1570 ttggacccccgtgtcctcctgcagaggaactctgtcagacccctgtgtgagcattccgga 1629
 DB 1501 TTGGACCCCGTGTCTCCACAGGAGACTCTGTCTGCGGACACTGTGTGAGCATTTGGA 1560
 QY 1630 aagaatttctcagactagttgaatgtcaatgttgcagactgttcagtgttcacatgataca 1689
 DB 1561 AAGATTTTCCGCTCCAGTTTATGATTAATTCAGATTGTGATGATGATGATGATGATGATGAT 1620
 QY 1690 atccatcctcgtgagcttccatgaatgatccagctgtccagcagcagcccccggttccct 1749
 DB 1621 ATTTCATCTGTGGAGCTTCTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1750 tctcgaacatacactacactcctcagataaataaaccatacactacactacactacact 1809
 DB 1681 TCTCGGACATGACACTACATCTCCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
 QY 1810 ggaaccatataaggtggtatgataatgataatgataatgataatgataatgataatgataat 1869
 DB 1741 GGAATCTGTTAATGTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1791
 QY 1870 aacaaatcaactactcagctcagcttccctg--gactagccagagagcagagcttgcagact 1927
 DB 1792 AACATTAACAACTCTTACCGGATTCCTCGAGCGATGAGCGAGAGAGAGCGCTTTGAGACT 1851
 QY 1928 cctgttggagacagctgtgtcagctcagctcagcagcagcagcagcagcagcagcagcagc 1987
 DB 1852 CCTGTGGGAGACACAGCTCGCTCAGCAG--CCGACCAAGAGAGCGCTGCTCGGC---ACCGGCT 1907
 QY 1988 gcttaagtgctgctcagagatgtcttatacaatgttaagatgagatgagatgagatgagat 2047
 DB 1908 GCTTAAGTGCTGCTCAGAGATGTCTT--TATCTTGTGTGAATGATGAGATTTCCAG 1966
 QY 2048 cctccctcc 2057
 DB 1967 CCTCCCTCC 1976

RESULT 9
 AF099932 1712 bp mRNA linear ROD 21-JAN-1999
 LOCUS AF099932
 DEFINITION Mus musculus beta-Trop protein E3RS-Ikappab mRNA, complete cds.
 ACCESSION AF099932
 VERSION AF099932.1 GI:4008019
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1712)
 AUTHORS Varon, A., Hatzuba, A., Davis, M., Lavon, I., Amit, S., Manning, A.M.,
 Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
 TITLE Identification of the receptor component of the
 JOURNAL Nature 396 (6711), 590-594 (1998)
 MEDLINE 99075339
 REFERENCE 2 (bases 1 to 1712)
 AUTHORS Varon, A., Hatzuba, A., Mercurio, F., Manning, A.M., Andersen, J.S.,
 Mann, M. and Ben-Neriah, Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem,
 Bin Karen, Jerusalem 91120, Israel
 FEATURES
 source
 CDS
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1..1710
 /note="beta-transducin repeat-containing protein; receptor
 component of Ikba-ubiquitin ligase"
 /product="beta-Trop protein E3RS-Ikappab"
 /protein_id="A008701.1"
 /db_xref="GI:4008020"
 /translation="MDPAEAVLQERALKFMNSEREDCNCEPRKIIPEKNSLRQY
 NSCARCINQDENVCLSTAMKTEENCVAKRLNGSSMSIVPQRKLSAYEKELCY
 KYEOWSESDIYEPVEHLISOMCHYOHGINSYLPMLORDITLTPARGIDHAENI
 LSYLAKSICAEELVCKEYRYTSGMIMKLIEMVNTDSIMRLARRRGQYLF
 KRPEENAPNSFRALYKTIIDILETBSMKGRSHSLRHCSESKVCLQYD
 DKIVSGLDNFTIKIWDKSTLECKRIILGTHSVLCIYDERVITGSSDSTVRWD
 NGENMLNTLIHCEAVLIHLRFNNGMAYVCSKDRSLAVMDMSPDITLRVYGHRA
 VVWVEPDDRYIVASASDRTIKVNTSTCEFTVTLNGHKGIALQYRDLVYSGSDN
 TLRMDIECGACLRVLEGHLELVRICRFENKRIIVSGAVDGRIKWDLMALIDPRAPAG
 TICTRLIVSHSRVRLQDEDEQYIVSSHDITILMDPLNDPMAIAEPRSPKTYT
 ISR"

BASE COUNT 469 a 399 c 453 g 391 t
 ORIGIN
 Query Match 67.5%; Score 1451.2; DB 10; Length 1712;
 Best Local Similarity 90.5%; Pred. No. 0;
 Matches 1549; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 70 atgagccggcggagcgggtgctgcagagaaagcactcaattatgaattcctcagag 129
 DB 1 ATGAGCCCGGCGGAGGCGGTGCTGCGAGAAAGCGCTTAAGTTATGATTCCTCAGAG 60
 QY 130 agagaagactgttaataatgagcgaacccctcagagaataatccagagaagaactcact 189
 DB 61 AGAGAAGACTGTATATATGCGGAACCCCTAGGAGATTAATACAGAGAAGATTCACCT 120
 QY 190 agacagacatacaacagcgtgtgcagactcgtcttaacaaagaaacagatgtttaga 249
 DB 121 AGACAGACTTACAAACGCTGTGCGAGGCTTTGCATTAACCAAGACACATATGTCTAAC 180
 QY 250 agcactgtctatgaagactgaattgtgtgcacaaacaaacaaactgtgcactcctcc 309
 DB 181 AGCAGTGTATGAAGACTGAAATGTGTGGCCAAAGCCAAACTGTGCCAATGCGACTTCC 240
 QY 310 agtatgtgtgtcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 369
 DB 241 AGCATATTTGTGCCCAAGAGCGGAAACCTCAGCAAGCTATGAGAAAGAAAGAGAGCTG 300
 QY 370 tgtgtcaaatctatgagcagtggtcagagtcagatcagatgagatgttgaacactct 429
 DB 301 TGTGTCAAGATTTTGAAGAGAGTGCTGATGATCAAGTGAATTTGTAGAACACCTT 360
 QY 430 atatcccaaatgtgtcattacacaaatggcacaataaactcgtatcttaacactatgttg 489
 DB 361 ATATCCCAATGTGTCACTACAGCATGGGACATCAATCACTTACTTAACATCTATGCTG 420
 QY 490 cagaagatattcactcgtctgcagcagcagcagcagcagcagcagcagcagcagcagcagc 549

```

Db 421 CAGAGGATTTATACAGTCTGCTGAGGAGGCTGAGCAGCATGCTGAGACATT 480
Qy 550 ctgtatacttgatgacaaatcaatcatatgtgtctgacatgtgtgcaagaatgttac 609
Db 481 CTGTATATTGAGAGCCAAAGTCACTGTGTGCTGTGAGCTGTGTGCAAGAAATGCTAC 540
Qy 610 ggaatgacctctgacatgtgtgtgtgaaagaagcttatcgagaagaatgttgagaagat 669
Db 541 CCGGAGACCTCGAGGAGGATGCTGTGAAAAAGCTCATCGAGAGGATGTGAGGAGC 600
Qy 670 tctctgtgaaagagcctgcaagaagaagatgtggagagatgtttttaaagaaga 729
Db 601 TCTCTGTGAGGAGGCTGTGAGAGGCTGAGAGGCTGTGAGAGGCTGTGAGAGGCT 660
Qy 730 cctctgacgaggaatgtctctcccaactcttttlaagaagacttaccataatata 789
Db 661 CCTCGTGAAGAGAGGCTCTCCCAAGCTCTTTATAGAGGCTTTATCTTAATAATCATTA 720
Qy 790 caagacattgagacaatgaatctaatgtgagagatgtgaaagaatgttttacaagaatt 849
Db 721 CAAGACATTGAGACAAATGAGTCAATTGAGATGTGGGCGCATAGTTTACAGAGAAATC 780
Qy 850 cacttcgaaagtgaaacaagaagagtttactgttttactaatgatatgataagaata 909
Db 781 CACTGCGGAGAGTGAACAAGTAAAGGCTTTACTTTTACATGACGACGACGACAGATTA 840
Qy 910 gtaagcgcctctgagacacacaatcaagatctggaataaacaatltggaatgtcaag 969
Db 841 GTCACGCGGCTTCTGAGACAAACACCTTAAGATCTGGATTAAGACACACTGGATGCAAG 900
Qy 970 gcaattctacagagcctacaaaggttcagtcctgtctcccgatgataagagatgtctc 1029
Db 901 CGGATTTCTACGCGGACACAGGCTCCGCTGCTGTGCAATGACGATGAGAGGCTGATC 960
Qy 1030 ataaacagatcatcagatccacaggtctgagatgtggaatgttaataacaggtgaaatgtca 1089
Db 961 ATCAGAGGCTCTCTGACAGCTCCACGCTGACAGAGTGTGGAATGAAACGAGGTGAGATGCTA 1020
Qy 1090 aacaagttgattcaacacatgtgaaagaagttctgcaattgcgtttcaataatgtgcatgtg 1149
Db 1021 AACCATTTGATTACACACACTGTGAAACCGCTTCTGACACGCTGCTTCAATATGCGCATGATG 1080
Qy 1150 gtagactgtctcaagaatcgttccatctgtctgtatgagatatgctcccaactacatt 1209
Db 1081 GTGACCTGTTCCAAAGACCGTGTCAATCGTGTGGAATGTGGCTTCCCACTGACATC 1140
Qy 1210 accctcggagaggtgtgtgtctgacacagcagctgtctcaatgtgtgactgtatgac 1269
Db 1141 ACCCTCAGAGGAGGCTGCTGTGAGACACCGAGCTGCGGTCAATGTTGTAGACTTTGATGAC 1200
Qy 1270 aagtaacattgttctgtcatctgtgagatagaactataaagtatggaacacaagtttgt 1329
Db 1201 AAGTACATCTGTTTCTGCTCTGAGATGAGAACCAATGAAGGTGTGAAACACAAGTACCTGT 1260
Qy 1330 gaatttgttaagaaccttaatgtgacacaaagagcaltgtctgtttgcatgacaggaac 1389
Db 1261 GAATTGTGAAGACCTTAATGTGGCAACAGCTGCGCATCGCTGTTTACAGTACGAGAAC 1320
Qy 1390 aagctgtgtagtggctgtcatctgacacaaatcatcagatgtgagatgtgtgt 1449
Db 1321 AGGCTGTGTGTAGGCGCTCTCTGTACAAACACATCAGGCTGTGAGACATGAGTGTGA 1380
Qy 1450 gcaatgttactagagtgltagaagagccatagaagaatgtgtgtgttactgattgataac 1509
Db 1381 GCATCTTGTGAGAGGCTGTGAGAGGCGCATGAGAGGTTGTGATGCTGATTTGATTAAC 1440
Qy 1510 aagaagatagtgatgtgggagcctatgataagaaatlaaagtgtgagatcctgtgtgtct 1569
Db 1441 AAAAGGATGATGAGGCGGAGCTTATATGGAATAATTAAGTGTGAGATCTTATGCTGCT 1500
Qy 1570 ttggagcccgctgtccctcggagagacactctgttcaagacccctgtgtgagacattcga 1629
Db 1501 TTGGACCCCGGCTGTGAGAGGAGGACTCTGTCTGTGCGGACACTTGTGTGAGCATTTCTGA 1560

```

```

Qy 1630 agagtttctgactacagttgattgaaatccagatgtgcatgattacatgatacaca 1689
Db 1561 AGAGTTTCCGCTCCAGTTGTGATGAAATTCAGATGTGTGAGTGTACATGATGACACA 1620
Qy 1690 atctcatctggagacttcccaatgatcatcagctgtgcccaagctgaaccccccttccct 1749
Db 1621 ATTTCATCTGTGGACTTCCGAATGATCCAGCTGCTCAGCTGAACCGCCCTCCCTCC 1680
Qy 1750 tctgacacatcacctacatctccagataat 1781
Db 1681 TCTGGACATTCACCTTCAATCTCCAGATAAT 1712

RESULT 10
AF112979 1710 bp mRNA linear ROD 02-MAR-1999
LOCUS
DEFINITION Mus musculus beta-transducin repeat containing protein mRNA,
complete cds.
ACCESSION AF112979
VERSION AF112979.1 GI:4140717
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1710)
Spencer, E., Jiang, J., and Chen, Z.J.
Signal-induced ubiquitination of IkappaBalpha by the F-box protein
Slimb/beta-Trcp
Genes Dev. 13 (3), 284-294 (1999)
99145465
MEDLINE
2 (bases 1 to 1710)
Chen, Z.J.
Direct Submission
Submitted (10-DEC-1998) Molecular Biology and Oncology, UT
Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX
75235-9148, USA
JOURNAL
TITLE
AUTHORS
REFERENCE
AUTHORS
CDS
1..1710
/organism="Mus musculus"
/db_xref="taxon:10090"
1..1710
/function="ubiquitinates phosphorylated IkbA"
/note="beta-trcp; F-box protein; Ikb-ubiquitin ligase;
substrate recognition subunit of SCF complex; similar to
Homo sapiens beta-Trcp and Drosophila melanogaster Slimb"
/codon_start=1
/product="beta-transducin repeat containing protein"
/protein_id="A0D04181.1"
/db_xref="GI:4140718"
/translation="MDPAEVLQEKALFKNSSEREDCNNGPPRKIIPEKNSLRQTY
NSCARLCTNOETVCLSTAMKTENCGYAKKLLNGSSMIVPQRLKLSAYEKEKELCV
KYFEQWSESDQVEFEVHLISOMCHYOHGILNSYIKMLODRDFTLAPAGDHLIENI
LSYLDKSLSCAATLVCEKEMRYTSDGMLKKLIERVRYRSDLSMRGLAERGRGQULFK
NRPDENAPNPSFYRALYPRITIDETTESNMKCGHSQRIHCHSETSKGYCLQYD
DOKIVSGLRNTIKIKMDSTLECKRLTHTOSVLCLOLGEERYITGSSDSTVRYADV
NAGEMLNTLIHCEAVLHLFPNMGMYTCSKDSIAYVMDASPTDITLRYLVGRHAA
VANVDDDKYIVASGDRTIKYVNTSTCEFRVNLNKKGIACLOYRDLRLVYSSDN
TIRLMDIEGACIRVLEGHDELVRCIRPNKRITVSGAYGCKIKVMDLMAADLPAPAG
TICLRILVHSGVFRLOPDEFQIYSSSHDDTILLPDLNPAHAHEPPRSKRYTY
ISR"
BASE COUNT 469 a 399 c 453 g 389 t
ORIGIN
Query Match 67.18; Score 1442.8; DB 10; Length 1710;
Best Local Similarity 90.2%; Prid. No. 0;
Matches 1543; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
Qy 70 atggaccggcgagggagtgatgtcgaagaagagcaatgatttgaattcctcaag 129
Db 1 ATGAGCCCGGAGAGGCGGTGCTGACAGAGAAAGCGCTTAAGTTTATGATTCCTCAGAG 60

```



```

/protein_id="AA02810.1"
/db_xref="GI:295543"
/translation="MEGFSCLQPTASEREDCNRPDKPIITEKNTLRQTKLANGT
SSMIVPQRLKISANEKEKEICLVYFEQMSDCDVEVEHLISMCHQOHMTYIK
PMLORDEITAPLARGDLHIAENILSYLDASLCSAEYLCKEWMYVTSQHMKILIER
MVRTSLMRGLAERNGOYLFKNRPPDGKTPKSPFYALPKITIDDETIESMRG
RHSIORHCHSESTKGYCLOYDOKITVSGIRPDNTIKIMKNTLECKRYVMGHGSLV
CLODEVITITGSSDSTVRVMDVTGEMNTLTHCEAVLHLRNNGMATCSKDRSI
AVMDASATIDTLERVLVGHRAAVNVDFDKYIVSASGDTIKVWNTSTSEFVTLN
GHRGIACLOYRDLVYSGSSDNTIRLMDLECGALRYLBEHELVIRIRPDNKRIVS
GAYDGKIKVMDLVVALDPRAPAGTILRTIIVHSGRVRFLQFDEFQIVSSHDTILI
WDFLNDPGLA"
1671
polya_site
/gene="beta-TRCP"
BASE COUNT 495 a 341 c 424 g 411 t
ORIGIN
Query Match 47.3%; Score 1017.8; DB 5; Length 1671;
Best Local Similarity 81.8%; Pred. No. 7.9e-282;
Matches 1175; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
QY 284 aacaaacttgcgaatgacatccatccatgattgtgcccgaacgaacgaatcctcaag 343
D 157 AGACAAACTTGCCTTACGCAATGCACTTCCACATGATGTGCGCCAGAGCAAACTGTGAG 216
QY 344 caagctatgaagaagaagaactgtgtcgaacttgcagcagtggttcagagtcag 403
D 217 CAATTTAGCAAGAGGAAAAGAGCTATGCTCAATATTTTGCAGCTGTCGCGAGTGCG 276
QY 404 atcaagtgaaattgttgaaacatctatcccaatgtgcatcacaacatgagcaca 463
D 277 ATCAGTAGAGATTGTTTAAACACCTGATATCTCGAATGTGCGACATATGACATGACATA 336
QY 464 taactcgtatctaaactatgttcgagagagatctcataactgtctgcagctcgg 523
D 337 TAAACACTTACCTTAAAGCANTGTACAAAGGAGATTTCATTACCGCACTACAGCTCGG 396
QY 524 gattgcatcatcgctgagaaacatctgtcatcactgagtgagcaactcaactatgctg 583
D 397 GACCTGATGACATAGCAGAAAATATACTTTTCACTCGATGCAAAAGTCAATGTTCTTG 456
QY 584 ctgaactgtgtgaaagaatgtaacgagactcgtatgagctgtgtgtgaaagaagc 643
D 457 CAGAACTGTATGTAGAGAGGTGTATCGAGTACCTCAGATGGAATGCTGTGGAAGAAC 516
QY 644 ttatcagaagaatgctgcagacagatctctgtgagagagcctgcagacgaacgaagat 703
D 517 TCATAGAGCGGATGCTCGCGACAGATTCTCTTTGAGAGAGAGCTGCGACAGAAAGAGAT 576
QY 704 ggggaacaglatlatacaaaaaaactcctgcagcgggaatgtctccccaactcttlt 763
D 577 GGGGCAATATTTTGTAAAAAACAACCTCCAGATGAGAAAAACCCACCAATTTATCTCT 636
QY 764 atagaacacttatcctaataatltacagaacatgagacatagaaatctaattgagat 823
D 637 ACAGAGCCCTTTACCCAAAATTTATTAAGACATAGACACATCAAGTCAACTGCGCT 696
QY 824 gtggaagacatagtttacagaagaatctcctgcgaagtgaaacgaacgaagagcttact 883
D 697 GTGGAGACACAGCTTACAAAGAAATTCCTGCGGAGTGAAGAAACAGCAAGGAGTATACT 756
QY 884 gttcaacgatgatgatcagaanaatagtaacgagcctcgcagagacaacaacatcaatc 943
D 757 GTCGCACTAGATGATGAGAGATAGTAGTAGTGCATCAGAGATTAACCACTTAAGATCT 816
QY 944 gggataaacaacatlggaatgaaaggaattctacagagccatcaagttcagtcctt 1003
D 817 GGGATTAAGAAATCTTTGGAGTGCAAGCGAGTGCTGATGGTCACTGGGTAGTTCTCT 876
QY 1004 gtctcagatgatgatgagatgtatcaacaagatcatcgatccacggtcagaagtc 1063
D 877 GTCGCAATATGATGAGAGATTAATCACTTACTGCTCTTCAGACTCTACCGTCCGGGTGT 936

```

```

QY 1064 gggatgtaataacaggtgaaatgtcctaacaagcttgatcaccatgttgaaagctgtc 1123
D 937 GCGAGCTGAACACAGAGAGAAATGTTGAACACACCTGATTCACCACTGTGAGCTGTCTC 996
QY 1124 acttcggttcaataatgacgatgatgagctgtctccaaatatggttcatctgtat 1183
D 997 ACTTGAGGTTTAAATATGATGATGATGCTACCTGCTCCAAACATGCTTCATTGACGTTT 1056
QY 1184 gggatagtcctccccaactgacatcaccctccgagaggtgtgtgctgcagaccagctg 1243
D 1057 GGGACATGCGCTCTGCGACCATATCACAATTACGAAGATTTCTGTGAGGCCACACGACTG 1116
QY 1244 ctgtcaatgtgtgacatcttgatgaaagatcatgtgttctgtcatctcgggagtagacta 1303
D 1117 CCGTAAATGCTGTGACTTGTGATGACAAATATATGCTCTGCATCTGCTGATGTCGAACAA 1176
QY 1304 taaagtgatgagacacaagaatgttgaaattgttgaaagccttaattgaaacaaagag 1363
D 1177 TAAAGTTTGAACACCACTACATGTAATGTTGTGCGGACATTTGAACGGCCACAAAGCTG 1236
QY 1364 gactgtcctgttgcagtaacagagcagctgtgtagtggtgtcatctgacaacacta 1423
D 1237 GTATTGACATGCTTACAGATTCAGATCGGCTGTGTGAGTGTGTTCTTTCGAACACAA 1296
QY 1424 tcagattatgagacataagaaatgtgtgtcatgtttacagatgttgaagagccatgaaat 1483
D 1297 TCAGATTGTGGGATATTGTAATCGGTCGATGTTTGGCGGTTCTGMAAGGACATGAAGAAC 1356
QY 1484 tgggtgtgtatctgattgataaagaagaatgtagtgcagtgagcctatgattgaaana 1543
D 1357 TGGTCCGGTGCAATCCGCTTTGATTAACAGAGAAATGTCGTGAGCANTATGACGGAAAAA 1416
QY 1544 taaagtgtgagatctgtgtgtctgttgagccctgtctcctgcagaggaacatctgtc 1603
D 1417 TAAAGTGTGGGACCTTGTGCGCTGTTGAGCCCGCTGCGCTGAGGAGACCTGTGTGTC 1476
QY 1604 taaggaacctgtgtgagaaatctccgaaagaagtttccagactacagtttgatccaga 1663
D 1477 TCAGGACTCTTGTGGAGCATTCGCGCAGAGATTTTCCGCTCGAGTTTGACGAGTTCCAGA 1536
QY 1664 ttgtgagtagttcacatgtagtgaacatctcatctcattggagacttccaaatgtagcag 1720
D 1537 TAGTAAGCAGCTCTCAAGATGACACTATCTCATCTGAGGATTTTTCACACGACCAG 1593

```

RESULT 12
AB033279
LOCUS Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
DEFINITION beta-TRCP2 isoform A, complete cds.
ACCESSION AB033279
VERSION AB033279.1 GI:7209808
KEYWORDS BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform A.
SOURCE Homo sapiens fetal lung cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS Koike,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
 Katoh,M.
TITLE Molecular cloning and genomic structure of the betaTRCP2 gene on
 chromosome 5q35.1
JOURNAL Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
MEDLINE 20160458
REFERENCE 2 (bases 1 to 2134)
AUTHORS Katoh,M.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center,
 Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
 (E-mail:mkatoh@nc.ccr.go.jp, Tel:81-3-3542-2511(ex.4402),
 Fax:81-3-3541-2685)
FEATURES location/Qualifiers

| source | gene | CDS | BASE COUNT |
|---|------|-----|------------|
| 1. .2134 | | | 566 a |
| /organism="Homo sapiens" | | | 477 c |
| /db_xref="taxon:9606" | | | 515 g |
| /chromosome="5" | | | 576 t |
| /map="5q35.1" | | | |
| /tissue_type="fetal lung" | | | |
| 102. .1628 | | | |
| /gene="BTRCP2" | | | |
| 102. .1628 | | | |
| /gene="BTRCP2" | | | |
| /codon_start=1 | | | |
| /product="F-box and WD-repeats protein beta-TrCP2 isoform A" | | | |
| /protein_id="BAAB2329.1" | | | |
| /db_xref="GI:7209809" | | | |
| /translation="MEPDSVIEDKTIETILMISNGTSSVYSKRPSEGNQYOKEDLCIK YFDMSIDQVEVEFHEHLISRMCHQHGHNISYLPRMQDFITFLPEQGDHIAENIL SYDARSLCAELCYCKEMORYISGMGLMKLIERWYDITPLMKLSIRRMQDYLFPK RPMQDPNSFYRSLYPKYKTIODIEPIESNMGRGNINLORIOCRSNSKGYVCLOYDEK IISLRNRSIKIMDKTILECLKVLITGTCSVYLQYDERIVITGSSSVYRWAVQVNG EVLNTLLIHNEVAHLIRFESNGLWYCSKDRSIAWMDASATDILRLRYLNGHRAAVV VDEPDKYIHSAGSGDRITIKWSTSTCEYFVRLTNGHRRISACLOYRDLVLSGSSINTLR LMDIEGACALRVLEGHELVLCIFEDNKRITVSGAYDRIKWMDLQALDPRAPSTSLC LRTIVHSGRVFRLQFDEFLQIISSHDDTILLIMDFLAVPESAQNETSPSRITYYISR " | | | |

| | | | | |
|----------------------------|-------|--------------------|----------|-------------|
| Query Match | 41.1% | Score 883.4 | DB 9 | Length 2134 |
| Best Local Similarity | 74.8% | Pred. No. 4.5e-243 | | |
| Matches 1122; Conservative | 0 | Mismatches 371 | Indels 6 | Gaps 1 |

| | | | |
|----|-----|--|-----|
| OY | 298 | aatgcacttcagtatgatgttgccccaagaacggaacctcagcaagcatltaaaag | 357 |
| Dd | 153 | AATGGAACATCATCTGTGATCGTCCAGAAAGAAGGCCATCAGAAAGCAACTATCAAAA | 212 |
| OY | 358 | gaaaaggacagtgtgtttaaataactttgaagcagtgctgaagtcagatcaagltgaaattc | 417 |
| Dd | 213 | GAAAAAGCTTGTAATTAAATATTTTTTCACAGAGTGCTGTAATCAGATCAAGTGGAAATT | 272 |
| OY | 418 | gtggaacatctatatcccaaatgytcatcattaccaaca tggcacataaacctcytatct | 477 |
| Dd | 273 | GTGGACAATCTTATTTCACGAATGTCATTTATCAGATGAGACATTTAATCTTTACCTG | 332 |
| OY | 478 | aaacctatgttcagagagatltcaaacctgcctgcagctcgaggatlygatcatac | 537 |
| Dd | 333 | AAGCCCATGTGCAGCGGGACTTTATTACCCTTTACACAGCAAGCCTTGATCATCAATA | 392 |
| OY | 538 | gctgaacacatcttcgtacacctgabatbcacaatbaactatgtcgtcgaacttgtgc | 597 |
| Dd | 393 | GCAGAAACATTTCTTTGCTTACCTGGATGCCAGGTCGTGTTGTCAGCGAGAGCTGSTATGT | 452 |
| OY | 598 | aaggaatgttacccgaagtyaacctctgaatgacatgctgtgtgtgaagaaactltagagaatg | 657 |
| Dd | 453 | AAAGATATGGCAGCGAGTATCTCAGAAAGATGCTTTGCAAGAACGTCATTTGAACGATNG | 512 |
| OY | 658 | gtcaggacaagttctctgtgagagagcctctgcagagaacaagaagatlyggcagatatta | 717 |
| Dd | 513 | GTACGACTGTATCCCCTTATGGAAGAAGACTTTCCAGAAAMAAGAAGGTGGATCAGTACTCG | 572 |
| OY | 718 | tltcaaaaaaacactctctgcagcggaatgcctctcccacactctlttataagaccattat | 777 |
| Dd | 573 | TTTTAAAAACAACCCACAGATGCC-----CCTCCAAATTATTTTATAGGTCAATTATAC | 626 |
| OY | 778 | cctaataatlatacagaacatltagacaatalgaatcttaatltagagatgltagaagaatagt | 837 |
| Dd | 627 | CCAAGATTTATCCAGGATATAGAGACTATACAAATCTAACTGGCGTGTGGACACACACAC | 686 |
| OY | 838 | ttacagaataactcactgcgcgaagtgaacaacgaagagatttactgtttacagatbat | 897 |
| Dd | 687 | TTTCAGAGAGATTTCAGTGGCGCTCTGAAATATAGTAAAGGTGCTTACTGTTTACAGTAGCAT | 746 |

| | | | |
|---|------|---|------|
| Oy | 898 | gaticgaataatagtaaggcccttcgagacaaacaacaatcaagaatttgggtataaaaacca | 957 |
| Dd | 747 | GATGAAAAAATTATCAGTGGCCCTACGAAGATAATTCTATTAAAGATATGGGATTAACCACGC | 806 |
| Oy | 958 | ttagaatcgaacgcgaattctccaagccacaaagttcaagtccttcgtccagtatgat | 1017 |
| Dd | 807 | CTGCAATGTMTTGAAAGTGTTAAACAGCACACAGCGCTGTCCTCTGCTCGTAGATCAT | 866 |
| Oy | 1018 | gagagaatgatcataaagaagatcatcgattccaacggtcagaagtgttggaigttaaataca | 1077 |
| Dd | 867 | GAGCGTGTGATTTGTAACACGCGCTTTACAATTCCTACGGGAGAAGTGTGGAGTGTGAACAG | 926 |
| Oy | 1078 | ggttgaatgctaaacaagtttgtattaccattgttgaagaagatcttgcacttggttcaat | 1137 |
| Dd | 927 | GGTGAAGTCTTTAACACATTGATCCACCACAATAGAGCTGTATTGTGACTTTACGCTTCAGC | 986 |
| Oy | 1138 | aatgcatgatgtgtagctctcctcaaagaatcglttccattctgtatggagatatggctccc | 1197 |
| Dd | 987 | AATGACATGATGTMACCTGTTCCAAAGACGCGCTCCATTGCTGTGTGGACATGGCTTCT | 1046 |
| Oy | 1198 | ccaactgacattaccctccgagaagtgctgttgcgaaccgcagctgcgtctaagtftta | 1257 |
| Dd | 1047 | GGACCGCAATCATTTCACGCCGTCTCTGGTGGCCACCGGGCTGCCGTCAATTAFTA | 1106 |
| Oy | 1258 | gaccttgaatgaatataattgtttcttcgatalcattgggagatagaactaaagttatggac | 1317 |
| Dd | 1107 | GACTTTGACACAAATACATGCTGTCTGCTCCTGTGTGACAGGACCATCAAGTCTGAGAC | 1166 |
| Oy | 1318 | acaagtaacttgtgaatttgttaaggaaccttaaatgtagcaacaagaagcattgctgtttg | 1377 |
| Dd | 1167 | ACGACACACTTGTAATTTGTTCTGTACTCTCATATGGGCAACAGCGGGCATTTGCCGTCTC | 1226 |
| Oy | 1378 | cagtaacaggaacagcgttgtagtgaatgagtcattcatctgacaacatacagattaaggac | 1437 |
| Dd | 1227 | CAGTACAGGGATCGCTCTGTTGTTAGTGGATCATCAGATAATFACATTBAGCTCTGGAT | 1286 |
| Oy | 1438 | atagaatgtgtgtcacgttttacgaagtgtttaagaagccatlgaynaatttgytctgtat | 1497 |
| Dd | 1287 | ATTGAATGTGGTGCCTGTTTAAAGTCTCTAGAGGACATGAAGAAATTTGGTCCGATGCATC | 1346 |
| Oy | 1498 | cgatttgtataacaagaagatagtcagtgaggccctatgtgtgaaaaataaagtgtggat | 1557 |
| Dd | 1347 | CGTTTTGTATACAAAGAGATTGTCTAGTGGGGCTATATGAGAAAATTAAAGTTTGGGAC | 1406 |
| Oy | 1558 | cttgytgatgcttbgagccccgtyctctccgcaggyaacactlgtctcaggaaccttgy | 1617 |
| Dd | 1407 | TTCGAAAGCTGCTTGACCCCTCGAACCCACAGAACACAACTGTGTTTCCGCACATTTGGTG | 1466 |
| Oy | 1618 | gagcatcccggaagaagtltttcogactacagtttgatgaatccagaatgtcagtagttca | 1677 |
| Dd | 1467 | GAACATTTGTGACGCTGTCTCGCTCCAGTTTGATGTTCAGATCATCAGCAGCTCC | 1526 |
| Oy | 1678 | catgtatgacaacaatccatcatggacttccctaaatgltcaagctgcgccaaagctgaacc | 1737 |
| Dd | 1527 | CATGATGTACACTATTATTATTTGGGATTTCTTAAATGTGCTCCACAGTCCACAGATGAG | 1586 |
| Oy | 1738 | cccccttcccccttcgcgacatacaaccatactccagataataaaccatacactgacc | 1796 |
| Dd | 1587 | ACCGTTCTCCCTCCAGAACATACACTTAACTCTCTAATATAAGAGTCTGCACATTTCACC | 1645 |
| RESULT 13 | | | |
| AB033280 2252 bp mRNA linear PRI 14-MAR-2000 | | | |
| LOCUS Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein | | | |
| DEFINITION beta-TRPC2 isoform B, complete cds. | | | |
| ACCESSION AB033280 | | | |
| KEYWORDS AB033280.1 GI:7209810 | | | |
| VERSION BTRCP2; F-box and WD-repeats protein beta-TRPC2 isoform B. | | | |
| SOURCE Homo sapiens fetal lung cDNA to mRNA. | | | |
| ORANISM Homo sapiens | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |

REFERENCE 1 (sites)
 AUTHORS Kojke, J., Sagara, N., Kirikoshi, H., Takagi, A., Miwa, T., Hirai, M. and
 TITLE Molecular cloning and genomic structure of the betaTRCP2 gene on
 chromosome 5q35.1
 JOURNAL Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
 MEDLINE 20160458
 REFERENCE 2 (bases 1 to 2252)
 AUTHORS Kato, M.
 TITLE Direct Submission
 JOURNAL Submitted (05-OCT-1999) Masaru Kato, National Cancer Center,
 Genetics Division; Tsukiji 5-chome, Chuo-Ku, Tokyo 104-0045, Japan
 (E-mail: m.kato@ncc.go.jp, Tel: 81-3-3542-2511 (ex. 4402),
 Fax: 81-3-3541-2685)

FEATURES
 source location/Qualifiers
 1. 2252
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q35.1"
 /tissue-type="fetal lung"
 157..1746
 /gene="TRCP2"
 157..1746
 /gene="TRCP2"
 /codon_start=1
 /product="F-box and WD-repeats protein beta-TRCP2 isoform
 B"
 /protein_id="BA02330.1"
 /db_xref="GI:7209611"
 translation="MEPDSVIEDKTIELMNTSVEDONEDSPKKNLIMQISNTSSV
 IISRRPSEGNVQKEDKIKYFDWSSDVEVEHLISMCYHQHINSYKPL
 ORDFTALPEQGLDIAENILSYDARSLCAEYLCKEMQVISEGMLMKLIERMV
 TDLKMGKSERGMDOYLFKNRPDPGPSYRSLYPIKIDIEIESNMCGHNIQ
 RLOCSSEKSKYVCLQYDDEKILISLRNLSIKIMDKTSLKCLVLTGHTGSLQYD
 ERYVITGSSDSVRYMVDVNTGEVNLTIHNEAVLHILPEENGLWYCSKHSIAVMD
 ASADITLRLVLRRAVNVDFDCKITVASGDRTRKWNSTCEPVRILNKHRC
 IACIOYRLVVGSSDNTIRLMDIEGACILVEHSELVRCILFMDKRLVSGAIDN
 KIKVMDLQALDPRAPASTLCRLTLEHSGRFLQFDEFOIISSHDDTLIMDFLN
 VPSAONETRSPSRITYISR"

BASE COUNT 598 a 501 c 558 g 595 t
 ORIGIN

Query Match 41.1%; Score 883.4; DB 9; Length 2252;
 Best Local Similarity 74.8%; Pred. No. 4.5e-243;
 Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

Db 631 GTACGACCTGATCCCTATGGAAGACCTTTCGAAGAAGAGGCTGATACGACCTG 690
 Qy 718 ttcataaacacacccctctgaaggaaatgctctcccaactctttatagaagcaattat 777
 Db 691 TTTAAAACAGACCCACAGATGGC-----CTCCAAATTCATTTATATGATCATATAC 744
 Qy 778 cctaaatatataagaacatltgagacaatagaacttaatttgagatgtggaagactagt 837
 Db 745 CCAAGATTTATCCAGATATAGAGACTATGAAATCTAACCTGAGGGGTGGACACACAAC 804
 Qy 838 ttacaagaaattacatgcgaagtgaagaacaagaaggaattactgctttacagatgat 897
 Db 805 TTGCAGAGGATTCAGTCCGCTCTGANAATATGTAAGGTCTCTACTGTTTACAGTACGAT 864
 Qy 898 gatcagaataatgaacgagcctctgcagacaacaacataagaatctgggataaacaaca 957
 Db 865 GATGAAAAAATTTATCATTGCGCTTACAGAGATTAATCTATTAAGATATGGGATTAACAC 924
 Qy 958 ttggaatgcgaagcaatctccacagccatacagtgctcaagtcctctgctccagatgat 1017
 Db 925 CTGGAATGTTTGAAGGTGTAAACGACACAGAGGCTCTGCTCTGTCTGCAATGATGAT 984
 Qy 1018 gaagagatgatcataaagaacatcatcgagattccacggtccagatgtgtggaatataca 1077
 Db 985 GAGGCTGATTTGATTAAGTGGCTCTTCAAGATTTCTACGCTGAGAGTGGATGGATGACAG 1044
 Qy 1078 ggtgaatgtctaaacacggtltgattaccatgttgaagcaatcttgcaacttgcgttcaat 1137
 Db 1045 GGTGAAGTTTCTTAACACATTGATCCACCAATGAGGCTGTATGTCATTACGCTTACG 1104
 Qy 1138 aatgcaatgatgtgagactgctccaaagatcgtctgctgctgtaaggaatagagctcc 1197
 Db 1105 AATGACATGATGGGACCTGTTCCAAAGACCGCTCCATTGCTGTGTGGACATGGCTTCT 1164
 Qy 1198 ccaactgacatcacccctccggaagggtgctggtcggacacgagctgctgtcaatgtgtga 1257
 Db 1165 GCGACCGACATCATTATACCGCGGTCTCTGTTGGCCACCGGCTGCGCTCAATAGTATGA 1224
 Qy 1258 gaacttgatgaacaaatgatactgttcttgcaactgagggatgaactataaagatgtgaac 1317
 Db 1225 GACTTTGACGACAACTATACATGCTGCTGCTGCTGCTGACGACGACATCAAGTGTGAGC 1284
 Qy 1318 acaagtactgtgaattgtgtgaaggacctaataatgacacaagaagcaatgtcgttctg 1377
 Db 1265 AGGACACCTGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1344
 Qy 1378 cagtaacaggaagcgtgtgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1437
 Db 1345 CAGTACAGGAGATCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1404
 Qy 1438 atagaatgtgtgcaatgtttacagagtggttagaaggccatlgaggaattggtgtgta 1497
 Db 1405 ATTGAATGGGCGCTGTTTAAAGTCTCTAGAGGACATGAAAGATTTGTCGATGATC 1464
 Qy 1498 cgatttgataaagaagataagtaagtggtggtggtggtggtggtggtggtggtggtggt 1557
 Db 1465 CGGTTTGATTAACAGAGATTTGATGAGGAGGCTGATGATGGAATAATTAAGTTGGGAC 1524
 Qy 1558 cttgtgagctgttggaaccccggtgtctctcgaggaagaacactgtgtctcggaacctgtg 1617
 Db 1525 TTGCAAGCTGCTCTTGCCTTGCAGCCCTGAGCCCGACACACATGTTGTTGCGCATTTGG 1584
 Qy 1618 gaggcaatccggaagagtttctgactaagtttgaatgaatccatgcaatgtcagtaattca 1677
 Db 1585 GAACATTTCTGGACGCTGTGTTGGCTCCAGTTTGAATGATTTCAATCATCAGCAGCTCC 1644
 Qy 1678 catgatgacaacatccatctcgtggaactcctaataatgatacgaagctgccaagctgaacc 1737
 Db 1645 CATGATGACATATTTTGAATTTGGGATTTCTTAATGTGCTCCACAGTCCCAAGATAG 1704
 Qy 1738 ccccgctccctctcgaacataacacactacatctcagtaataacatacctacagcagc 1796

PR 09-DEC-1998: 98FR-0015545.
 PR 30-JAN-1998: 98FR-0001100.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 XX Arenzana Seisdedos F, Benarous R, Concorde J, Durand H;
 PI Kroll M, Margolin F,
 XX
 DR MPI, 1999-469329/39.
 DR P-PSDB; AA124054.
 XX
 PT New human beta-transducin repeat containing protein and its
 fragments useful as, or to screen for, antiviral, antitumour,
 anti-inflammatory and anti-Alzheimer's agents
 XX
 PS Claim 7; Page 57-60; 71pp: French.
 XX
 CC The present sequence encodes a human beta-transducin repeat containing
 CC protein (beta-Trcp). The protein directs proteins to the proteosome
 CC degradation pathways. The protein is able to interact with the Ypu
 CC protein of human immune deficiency virus-1 (HIV-1), cellular proteins
 CC IkappaB or beta-catenin (bc) and/or protein Skrip. The protein controls
 CC ubiquitinylation of phosphorylated proteins and thus their targeting to
 CC proteosomes for degradation. Depending on whether the process is
 CC inhibited or promoted, the result may be delayed breakdown of CD4 (in
 CC cases of HIV-1 infection); increased activity of IKB (and thus reduced
 CC activity of NFkappaB) and increased degradation of mutant bc in tumour
 CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
 CC antitumour agents that disrupt cell cycle regulation or protein
 CC degradation in human tumour cells, and anti-inflammatory agents that
 CC disrupt activation by NFkappaB. Fragments of the protein are also
 CC useful for treating osteo-articular inflammation or acute inflammation
 CC associated with release of tumour necrosis factor.
 XX
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other:

Query Match 100.0%; Score 2151; DB 20; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgcctgacacaaaggcgcccgcgagagcgagaccagtgcc 60
 |||||
 Db 1 tgcgttgctgcgcctgacacaaaggcgcccgcgagagcgagaccagtgcc 60
 QY 61 tgcggcattatgacccgcccagggcggtgctgcaagagaagacccaattatgaat 120
 |||||
 Db 61 tgcggcattatgacccgcccagggcggtgctgcaagagaagacccaattatgaat 120
 QY 121 tccctcagagagaagactgttaataatggcgaacccctcagagaataatacacaagaag 180
 |||||
 Db 121 tccctcagagagaagactgttaataatggcgaacccctcagagaataatacacaagaag 180
 QY 181 aattcactaagcagacatacaacagcgtgctgcagacccctcctaaccagaacagta 240
 |||||
 Db 181 aattcactaagcagacatacaacagcgtgctgcagacccctcctaaccagaacagta 240
 QY 241 tctttagcaagacgtcctaagaagctgagaattgttgcccaaaaacttgcacat 300
 |||||
 Db 241 tctttagcaagacgtcctaagaagctgagaattgttgcccaaaaacttgcacat 300
 QY 301 ggcactccagatgatgtgtgcccagaacaggaactctcagcaagctatgaaagaagaa 360
 |||||
 Db 301 ggcactccagatgatgtgtgcccagaacaggaactctcagcaagctatgaaagaagaa 360
 QY 361 aaggaacgtgtgtcaataactttagcagcgtgtcagagtcagatcaagtggaattgtg 420
 |||||
 Db 361 aaggaacgtgtgtcaataactttagcagcgtgtcagagtcagatcaagtggaattgtg 420
 QY 421 gaaactatatacccaaatgtgtcattacaaatgagcacaatgagcacaataactgtatcttaaa 480
 |||||

Db 421 gaaactatatacccaaatgtgtcattacaaatgagcacaatgagcacaataactgtatcttaaa 480
 |||||
 QY 481 cctatgttcagagagatttcatcactgctgcgcagctcggggattggatcatcgtc 540
 |||||
 Db 481 cctatgttcagagagatttcatcactgctgcgcagctcggggattggatcatcgtc 540
 QY 541 gagaacattcgtcactacgtgatgccaaatcactatgtgtcgtcgaacttgtgtcaag 600
 |||||
 Db 541 gagaacattcgtcactacgtgatgccaaatcactatgtgtcgtcgaacttgtgtcaag 600
 QY 601 gaatgttcagagagatttcatcactgctgcgcagctcggggattggatcatcgtc 660
 |||||
 Db 601 gaatgttcagagagatttcatcactgctgcgcagctcggggattggatcatcgtc 660
 QY 661 aggaacagattcctcgtgtgagagagccttgccagaacgaagagatggaggacattatc 720
 |||||
 Db 661 aggaacagattcctcgtgtgagagagccttgccagaacgaagagatggaggacattatc 720
 QY 721 aaaaacaaactcctcgtgagagagccttgccagaacgaagagatggaggacattatc 780
 |||||
 Db 721 aaaaacaaactcctcgtgagagagccttgccagaacgaagagatggaggacattatc 780
 QY 781 aaaaattacagaagcattgagacacatagaatcctaatgtgagatgtggaacataagttta 840
 |||||
 Db 781 aaaaattacagaagcattgagacacatagaatcctaatgtgagatgtggaacataagttta 840
 QY 841 cagagaattcactgcgcgaagtgtgaaacgaacgaagagttactgtttacagtatgat 900
 |||||
 Db 841 cagagaattcactgcgcgaagtgtgaaacgaacgaagagttactgtttacagtatgat 900
 QY 901 cagaagaatagaaacggcgcttcgagacacaacacaaatcctcgggataaaacacatg 960
 |||||
 Db 901 cagaagaatagaaacggcgcttcgagacacaacacaaatcctcgggataaaacacatg 960
 QY 961 gaatgcagagcgaattcctcagaagcctacagatgagttcctcgtcctcagatgtgag 1020
 |||||
 Db 961 gaatgcagagcgaattcctcagaagcctacagatgagttcctcgtcctcagatgtgag 1020
 QY 1021 aagatgtcatatacagaagatcctcagatcctcagatcctcagatcctcagatcctcagat 1080
 |||||
 Db 1021 aagatgtcatatacagaagatcctcagatcctcagatcctcagatcctcagatcctcagat 1080
 QY 1081 gaaatgtcacaacagtttcatcaccattgtgaaagcagttcgtcacttgccttcaataat 1140
 |||||
 Db 1081 gaaatgtcacaacagtttcatcaccattgtgaaagcagttcgtcacttgccttcaataat 1140
 QY 1141 ggcattgtgtgacacctcctcaagaatcgttccattgtgtatgtatgtgcctccca 1200
 |||||
 Db 1141 ggcattgtgtgacacctcctcaagaatcgttccattgtgtatgtatgtgcctccca 1200
 QY 1201 actacacattacccctcggagaggtgctgctgcgacacacacgtcgtcattgtgtgagc 1260
 |||||
 Db 1201 actacacattacccctcggagaggtgctgctgcgacacacacgtcgtcattgtgtgagc 1260
 QY 1261 ttgtatgacaagaatcattgttctcgaatcctgggagatagaactaaaggtatgtgaacaca 1320
 |||||
 Db 1261 ttgtatgacaagaatcattgttctcgaatcctgggagatagaactaaaggtatgtgaacaca 1320
 QY 1321 agtaactgtgaattttgaaagaccttaaatgagcacaacgaagacatgcccgtttgcag 1380
 |||||
 Db 1321 agtaactgtgaattttgaaagaccttaaatgagcacaacgaagacatgcccgtttgcag 1380
 QY 1381 taacagggacagcgtctgtgtagtgcctcactcgaacacactcagatcatatggagacata 1440
 |||||
 Db 1381 taacagggacagcgtctgtgtagtgcctcactcgaacacactcagatcatatggagacata 1440
 QY 1441 gaatgtgtgacatgttctcagagtggttaagaagcgaatggaatgtgtcgtgtatcga 1500
 |||||
 Db 1441 gaatgtgtgacatgttctcagagtggttaagaagcgaatggaatgtgtcgtgtatcga 1500
 QY 1501 ttgtatacaagaagagatgtagtgggagcctatgtaggaataaagtgtggagctct 1560
 |||||

Db 1501 ttgataacaagaagatgctcagtgaggccatgatgtgaaaaaattaaagtgtggaattc 1560
 QY gtgagctgtcttgaccgccctgctcctcaggaagacactgtctacagaccctgtgag 1620
 Db 1561 gtgagctgtcttgaccgccctgctcctcaggaagacactgtctacagaccctgtgag 1620
 QY 1621 cactcggagaagatttctcagctacagtttgatgaattccagattgtcagtagttacat 1680
 Db 1621 cactcggagaagatttctcagctacagtttgatgaattccagattgtcagtagttacat 1680
 QY 1681 gatgacaacaatctcatcttgagctcttaaatgatacagctccaggaagaccccc 1740
 Db 1681 gatgacaacaatctcatcttgagctcttaaatgatacagctccaggaagaccccc 1740
 QY 1741 cgttcctctctcgaacatacaccatactccagataataaacatacactgacctat 1800
 Db 1741 cgttcctctctcgaacatacaccatactccagataataaacatacactgacctat 1800
 QY 1801 acttgccaggaaccataaagttcggtatttaacgtatctgccaatacagagatgagc 1860
 Db 1801 acttgccaggaaccataaagttcggtatttaacgtatctgccaatacagagatgagc 1860
 QY 1861 aacacaagaatcaatcaactactgcccagttccctcagctagccgaggaagcagagctt 1920
 Db 1861 aacacaagaatcaatcaactactgcccagttccctcagctagccgaggaagcagagctt 1920
 QY 1921 tgaagactcctgttggagacagttgtctcagctcgcgcagccaggaagcgtctactacagaca 1980
 Db 1921 tgaagactcctgttggagacagttgtctcagctcgcgcagccaggaagcgtctactacagaca 1980
 QY 1981 actgactgtctcagtgctgtatcagaagaatgtctctatcaatgtgaaatgttggaac 2040
 Db 1981 actgactgtctcagtgctgtatcagaagaatgtctctatcaatgtgaaatgttggaac 2040
 QY 2041 tttaaaccctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2100
 Db 2041 tttaaaccctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2100
 QY 2101 cagacaagaagtactataataattagttgttttcgcagaaaaa 2151
 Db 2101 cagacaagaagtactataataattagttgttttcgcagaaaaa 2151
 RESULT 2
 ID AAAS1229 standard; DNA; 2151 BP.
 AC AAAS1229;
 XX
 DT 26-SEP-2000 (first entry)
 DE Human beta-TrCP coding sequence.
 XX
 KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
 KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
 KW anti-inflammatory; immunosuppressive; cytosolic; ds.
 XX
 OS Homo sapiens.
 FH Key location/Qualifiers
 FT CDS 70..1779
 FT /tag= a
 FT /product= Human_TrCP
 PN WO200034447-A2.
 PD 15-JUN-2000.
 XX 10-DEC-1999; 99WO-US29371.
 PF 10-DEC-1998; 98US-0210060.
 PR
 XX (SIGN-) SIGNAL PHARM INC.
 PA

PA (YISS) YISSUM RES & DEV CO.
 XX Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
 PI Lavon I, Yaron A;
 XX
 DR WPI: 2000-431294/37.
 DR P-PSDB: AAY96697.
 XX
 PT Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
 PT treating disorder associated with NF-kappaB activation e.g. cancer,
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its
 PT variant
 XX
 PS Disclosure; Page 72; 77pp; English.
 XX

This DNA encodes human beta-TrCP, an F-box/WD protein family member,
 which has been shown to have homology to human E3 ubiquitin ligase (E3).
 E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor
 of protein of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
 degradation via the ubiquitin pathway is useful for identifying
 modulators of this process for use in treating diseases associated with
 activation of NF-kappa-B. In vitro analysis suggests that deletion of
 the F-box results in a protein that functions as a dominant negative
 molecule in vivo. Transient over-expression of delta-beta-TrCP (a
 deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
 in stimulated Jurkat cells, resulting in accumulation of phosphorylated
 I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
 activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
 inflammatory diseases, autoimmune diseases, cancer and viral infections.
 XX
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgcctggcaccacaaaggggggccgggagagggccacgtggcc 60
 Db 1 tgcgttgctgcgcctggcaccacaaaggggggccgggagagggccacgtggcc 60
 QY 61 tcggcagattatggaccggccggcggcgtgctgcaagagaagcaccacgttatgat 120
 Db 61 tcggcagattatggaccggccggcggcgtgctgcaagagaagcaccacgttatgat 120
 QY 121 tccctcagagagaagactgttaataatggcgaacccctagggaagataatcagagaag 180
 Db 121 tccctcagagagaagactgttaataatggcgaacccctagggaagataatcagagaag 180
 QY 181 aattcacttagacagacatacaacagctgtgcccagctctgtcttaaacccaagaacgta 240
 Db 181 aattcacttagacagacatacaacagctgtgcccagctctgtcttaaacccaagaacgta 240
 QY 241 ttttagaagacacgtgctatgaagactgtgctgtgcccacaaacaaactgtccaat 300
 Db 241 ttttagaagacacgtgctatgaagactgtgctgtgcccacaaacaaactgtccaat 300
 QY 301 ggcacttcagatagattgtgcccagcgaacactcagcaagctatgaagaagaa 360
 Db 301 ggcacttcagatagattgtgcccagcgaacactcagcaagctatgaagaagaa 360
 QY 361 aaggaactgtgtcaaatcttgaagcagtgctcagagtgatcaagtggaattgtg 420
 Db 361 aaggaactgtgtcaaatcttgaagcagtgctcagagtgatcaagtggaattgtg 420
 QY 421 gaacattatatacccaaatgtgcatcacaacatgggcacataaactctatcttaa 480
 Db 421 gaacattatatacccaaatgtgcatcacaacatgggcacataaactctatcttaa 480
 QY 481 cctatgttcagagaagatctcataactgtcctgcagctcggggatgtgatatcagct 540
 Db 481 cctatgttcagagaagatctcataactgtcctgcagctcggggatgtgatatcagct 540

| | | | |
|----|------|---|------|
| QY | 541 | gqgaacattctgtatactaccctggatgcacaaatcactatgctgtgtcgaactctgtgtgcaag | 600 |
| Db | 541 | gggaacattctgtatacacttgatgcacaaatcactatgctgtgtcgaactctgtgtgcaag | 600 |
| QY | 601 | gatatgtaccgaagtaccctcgtatgcatctctctgaaagaagcttctcgaagaatgctc | 660 |
| Db | 601 | gatatgtaccgaagtaccctcgtatgcatctctctgaaagaagcttctcgaagaatgctc | 660 |
| QY | 661 | aggaacattctctgtgtgagaagcctctgacgaacgaagaagatggtggagacattatctc | 720 |
| Db | 661 | aggaacattctctgtgtgagaagcctctgacgaacgaagaagatggtggagacattatctc | 720 |
| QY | 721 | aaaaacaaactctctgacggaaatgctcctcccaactctttatatgagacattatcct | 780 |
| Db | 721 | aaaaacaaactctctgacggaaatgctcctcccaactctttatatgagacattatcct | 780 |
| QY | 781 | aaaattatacagaacatttgagacaataagaaatctaatcttgagatgctgaaagacatagttca | 840 |
| Db | 781 | aaaattatacagaacatttgagacaataagaaatctaatcttgagatgctgaaagacatagttca | 840 |
| QY | 841 | cagaaatctacctcgcgaagtgtgaaacgaagaagggtttactgcttcaagtatgcat | 900 |
| Db | 841 | cagaaatctacctcgcgaagtgtgaaacgaagaagggtttactgcttcaagtatgcat | 900 |
| QY | 901 | cagaaatctacctcgcgaagtgtgaaacgaagaagggtttactgcttcaagtatgcat | 960 |
| Db | 901 | cagaaatctacctcgcgaagtgtgaaacgaagaagggtttactgcttcaagtatgcat | 960 |
| QY | 961 | gaatgcgaagcgaattctccacaagccatacaggtttcagctctctgtctccagatgatag | 1020 |
| Db | 961 | gaatgcgaagcgaattctccacaagccatacaggtttcagctctctgtctccagatgatag | 1020 |
| QY | 1021 | agagtgatcatataacaggatcatatcgagatccacagctgataagctgtgagatgtaataacagtt | 1080 |
| Db | 1021 | agagtgatcatataacaggatcatatcgagatccacagctgataagctgtgagatgtaataacagtt | 1080 |
| QY | 1081 | gaatgtcctaaacacagcttgatctccactcttggaagcagcttctcgaacttcgattcaatat | 1140 |
| Db | 1081 | gaatgtcctaaacacagcttgatctccactcttggaagcagcttctcgaacttcgattcaatat | 1140 |
| QY | 1141 | ggcatgtagtgcacctctgctccaaagatctgtccattgtcgtatgagatctggtcccccaca | 1200 |
| Db | 1141 | ggcatgtagtgcacctctgctccaaagatctgtccattgtcgtatgagatctggtcccccaca | 1200 |
| QY | 1201 | actgaacttaccctccggaaggtgtcgtgtccgcgaacacagcagctgcgtcaatgtgtgtagac | 1260 |
| Db | 1201 | actgaacttaccctccggaaggtgtcgtgtccgcgaacacagcagctgcgtcaatgtgtgtagac | 1260 |
| QY | 1261 | tttatacagaatgacattgttcttgatcatctcggagatagaaactataaagttatggaacaca | 1320 |
| Db | 1261 | tttatacagaatgacattgttcttgatcatctcggagatagaaactataaagttatggaacaca | 1320 |
| QY | 1321 | agtaactgtgaaatttgttaagacccttaaatgtgacacaaacgaagcattgcactgtgttgag | 1380 |
| Db | 1321 | agtaactgtgaaatttgttaagacccttaaatgtgacacaaacgaagcattgcactgtgttgag | 1380 |
| QY | 1381 | tacaggaacacggctgcgtgtaagtgtatgtgctatctgcacaaactataagttatggaacata | 1440 |
| Db | 1381 | tacaggaacacggctgcgtgtaagtgtatgtgctatctgcacaaactataagttatggaacata | 1440 |
| QY | 1441 | gaatgtgtgcatgttatacagagtgttataagaagccatagaaatctgtgctgttatcgaa | 1500 |
| Db | 1441 | gaatgtgtgcatgttatacagagtgttataagaagccatagaaatctgtgctgttatcgaa | 1500 |
| QY | 1501 | tttgataacaagaaggtatgtcagttggtgacctatgatatggaataataaagtgtggatcctt | 1560 |
| Db | 1501 | tttgataacaagaaggtatgtcagttggtgacctatgatatggaataataaagtgtggatcctt | 1560 |
| QY | 1561 | gttgagctcctcttgagaccccgctgcctctgacgaagaacacctctacaggaacccctgtgag | 1620 |
| Db | 1561 | gttgagctcctcttgagaccccgctgcctctgacgaagaacacctctctacaggaacccctgtgag | 1620 |
| QY | 1621 | catctccggaagagtttctgcactacaagtttgaattccagatgttcaagtagttacat | 1680 |

| | | | |
|---|------|--|------|
| Dd | 1621 | cattccggaagaagtttctgcactacagtcttgatgaattccaggctcgttcagtagttcacat | 1680 |
| Qy | 1681 | gatgacaacaatcccatcattggagcttctctaatagatacccgctcggccaagctgaaccccc | 1740 |
| Dd | 1681 | gatgacaacaatcccatcattggagcttctctaatagatacccgctcggccaagctgaaccccc | 1740 |
| Qy | 1741 | cgttcccctcttcgaaacatacaaccatacatctcaga.tlaaalaaacatacagacctcat | 1800 |
| Dd | 1741 | cgttcccctcttcgaaacatacaaccatacatctcaga.tlaaalaaacatacagacctcat | 1800 |
| Qy | 1801 | actgtcccaaggaccocattaaatgttggtgtttaaqtatatctgccaatatccagatgagc | 1860 |
| Dd | 1801 | actgtcccaaggaccocattaaatgttggtgtttaaqtatatctgccaatatccagatgagc | 1860 |
| Qy | 1861 | aacaacagtaaacaatacaactactgtcccaagtttcccttgactagccgagagcaagagctt | 1920 |
| Dd | 1861 | aacaacagtaaacaatacaactactgtcccaagtttcccttgactagccgagagcaagagctt | 1920 |
| Qy | 1921 | tgaagctccctcttgygacaacaagtgtgtctcgtcagtcgagcccgagagcgbtctactaagaca | 1980 |
| Dd | 1921 | tgaagctccctcttgygacaacaagtgtgtctcgtcagtcgagcccgagagcgbtctactaagaca | 1980 |
| Qy | 1981 | actgactccttcagtgctgcgtcatcagaaga.tgtctctcataltaatigtatgaattgagac | 2040 |
| Dd | 1981 | actgactccttcagtgctgcgtcatcagaaga.tgtctctcataltaatigtatgaattgagac | 2040 |
| Qy | 2041 | ttttaaacctccctcctctcctctccttcaacctcgtcacctagttctttcccatgtgtc | 2100 |
| Dd | 2041 | ttttaaacctccctcctcctcctccttcaacctcgtcacctagttctttcccatgtgtc | 2100 |
| Qy | 2101 | cagacaagaagtgacttaataatatattaagtgtttgtccagaaaaaaa 2151 | |
| Dd | 2101 | cagacaagaagtgacttaataatatattaagtgtttgtccagaaaaaaa 2151 | |
| RESULT 3 AAZ93350 ID AAZ93350 standard; cDNA; 2151 BP. XX AC AC AAZ93350; XX DfT 16-AUG-2000 (first entry) XX DE Sequence encoding F-box protein FBP-1. XX KW F-box protein; FBP; diagnosis; treatment; screening; agonist; KW antagonist; proliferative disorder; differentialtiave disorder; KW breast cancer; prostate cancer; ovarian cancer; cancer; KW small cell lung carcinoma; immune disorder; cardiovascular disorder; KW inflammatory disorder; human; ss. XX OS Homo sapiens. XX FH Key Location/Qualifiers FT CDS 70..1779 ET /*tag= a FT /product= FBP-1 XX FN MO200012679-Al. XX PD 09-MAR-2000.) XX PE 27-AUG-1999; 99MO-US19560. XX PR 28-AUG-1998; 98US-0098355. PR 03-FEB-1999; 99US-0118568. PR 15-MAR-1999; 99US-0124449. XX FA (UNITV) UNITV NEW YORK STATE. X1 Chlaur DS, Pagano M, Latres E; X2 | | | |

CC liases) which can be used for the targetted degradation of a target
CC polypeptide in vivo. Targetted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or
CC increasing the level of a target polypeptide and for creating and
CC expressing a destabilized polypeptide which is subjected to SCF
CC mediated proteolysis. Degrading any desired protein in a cell is
CC useful for preventing or treating diseases caused by the presence of
CC abnormal amount of the specific polypeptides, for drug discovery and
CC for gene therapy. Diseases treated include cancer, by degradation of
CC oncoproteins, Huntington's disease, other proliferative disorders and
CC microbial infections. The method provides a quick and easy
CC alternative to gene knockout technology. The target polypeptide can
CC be degraded at all stages, or a specific stage, of development in the
CC mature animal.
CC
CC
SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

1 tgccttgccgagcaccacaaaggggcgcccgcggaagagcgagcccaatgcc 60

50

Db 6

QV 12

Db 12.

QY 18:

Db 18:

QY 243

Db 243

QY 302

302 Db

36-207

20-

187

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

2

2
3
4

ON EAST

601 05

667 0v

Dh 667

0V 721

Query Match 100.0%; Score 2151; DB 22; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgctgcgcacccaagggcgcccgcgagagcgagccagctggcc 60
DB 1 tgcgttgctgcgctgcgcacccaagggcgcccgcgagagcgagccagctggcc 60
QY 61 tcggcgattatagacccggcgagggcggtgctgcgaagaaagcactcaattatgaat 120
DB 61 tcggcgattatagacccggcgagggcggtgctgcgaagaaagcactcaattatgaat 120
QY 121 tccctcagagagagagactgtaataatgycgaacccctcaggaataataaccagagaag 180
DB 121 tccctcagagagagagactgtaataatgycgaacccctcaggaataataaccagagaag 180
QY 181 aattcaacttagacagacatacaacagctgycgaactctgtcttaaccaagaacagta 240
DB 181 aattcaacttagacagacatacaacagctgycgaactctgtcttaaccaagaacagta 240
QY 241 tcttttagcaagacactgctcgtatgaagactgagaaattgtgycgaacaaacaaacttcacat 300
DB 241 tcttttagcaagacactgctcgtatgaagactgagaaattgtgycgaacaaacaaacttcacat 300
QY 301 ggcactccagatgattgtgcccgaagcaacgaaactctcagcaagctatgtaaaagaa 360
DB 301 ggcactccagatgattgtgcccgaagcaacgaaactctcagcaagctatgtaaaagaa 360
QY 361 aagggaactgtgttaaatattgagcagctgctcgaagctcagatcgaagtggaaattgtg 420
DB 361 aagggaactgtgttaaatattgagcagctgctcgaagctcagatcgaagtggaaattgtg 420
QY 421 gaacatctata tcccaaatgtgcatatcaacaatgycgaacataaacctgcatcttaa 480
DB 421 gaacatctata tcccaaatgtgcatatcaacaatgycgaacataaacctgcatcttaa 480
QY 481 cctatgttgcagagagatttcaatactgctctgcagagctcgaggatgtgcatatcgct 540
DB 481 cctatgttgcagagagatttcaatactgctctgcagagctcgaggatgtgcatatcgct 540
QY 541 gagaacattctgtactatgatacgaatcactatgtcgtcgaactgtgtgcaag 600
DB 541 gagaacattctgtactatgatacgaatcactatgtcgtcgaactgtgtgcaag 600
QY 601 gaattgtaccgagtgacctctgatgtgcacgtctgtgaaagaaactatcgagaaatgtgc 660
DB 601 gaattgtaccgagtgacctctgatgtgcacgtctgtgaaagaaactatcgagaaatgtgc 660
QY 661 aggaacagattctctgtgagagagcctgycgaagaaagagatgaggacaglatatttc 720
DB 661 aggaacagattctctgtgagagagcctgycgaagaaagagatgaggacaglatatttc 720
QY 721 aaaaacaacacctgcgagcggaatgctctcccaactcttttaagagcacttatcc 780
DB 721 aaaaacaacacctgcgagcggaatgctctcccaactcttttaagagcacttatcc 780
QY 781 aaaaattatacaagacattgagacaatagaaatctaattgagagatgtggaagacatagttta 840
DB 781 aaaaattatacaagacattgagacaatagaaatctaattgagagatgtggaagacatagttta 840
QY 841 cagagaattctcctgcgagagtgaaacaaagcaagagatttactgttttaacatagatgat 900
DB 841 cagagaattctcctgcgagagtgaaacaaagcaagagatttactgttttaacatagatgat 900
QY 901 cagaaaatagtaagcgagccttcgagacacacaacaaactcaagactcggatgataaacaacttg 960
DB 901 cagaaaatagtaagcgagccttcgagacacacaacaaactcaagactcggatgataaacaacttg 960
QY 961 gaatgcaaggaattctcaacagccataacaggttcagctcctctgtctccagatgagatgag 1020
DB 961 gaatgcaaggaattctcaacagccataacaggttcagctcctctgtctccagatgagatgag 1020
QY 1021 agagtgatcatacaagagatcagatccacggtcagagtggtgagatgtaataacagatc 1080

DB 1021 agagtgatcatacaagagatcagatccacggtcagagtggtgagatgtaataacagatc 1080
QY 1081 gaaatgtcaaacacagtttgatcaccattgtgaaagcagttctgcacactgctgtccaataat 1140
DB 1081 gaaatgtcaaacacagtttgatcaccattgtgaaagcagttctgcacactgctgtccaataat 1140
QY 1141 ggcattgattgtgacctgctccaagaatcgttccattgctgtatagggatatgacctccca 1200
DB 1141 ggcattgattgtgacctgctccaagaatcgttccattgctgtatagggatatgacctccca 1200
QY 1201 actgacattacccctcggagaggtgtgctgctgcggaacccgagctgctcgaattgtgtagac 1260
DB 1201 actgacattacccctcggagaggtgtgctgctgcggaacccgagctgctcgaattgtgtagac 1260
QY 1261 ttgtagacaagatcattgttctgcatactcgtgggatatagaactaataaggtatgaaacaca 1320
DB 1261 ttgtagacaagatcattgttctgcatactcgtgggatatagaactaataaggtatgaaacaca 1320
QY 1321 agtactgtgaaattgtgaaagacctaataatggaacaaacagagcaattgcctgttgcaag 1380
DB 1321 agtactgtgaaattgtgaaagacctaataatggaacaaacagagcaattgcctgttgcaag 1380
QY 1381 tacagagacagcgctgtagtagtgctcatctgcgaacacatacagattatagggacata 1440
DB 1381 tacagagacagcgctgtagtagtgctcatctgcgaacacatacagattatagggacata 1440
QY 1441 gaatgtgtgcatgatttaagagtgtagagtgtagagccatgaggaattgtgctgttatcga 1500
DB 1441 gaatgtgtgcatgatttaagagtgtagagtgtagagccatgaggaattgtgctgttatcga 1500
QY 1501 ttgtatatacaagagagatagctcagtgggcctcagatgagaaataaattgagggactct 1560
DB 1501 ttgtatatacaagagagatagctcagtgggcctcagatgagaaataaattgagggactct 1560
QY 1561 gtgctgtcttggaccccgctgctcctgcagggagacactgttlaacgagccctgtgagag 1620
DB 1561 gtgctgtcttggaccccgctgctcctgcagggagacactgttlaacgagccctgtgagag 1620
QY 1621 cattccggagaagatttttgcactacaagtttgatgaattccagattgcagatgttccat 1680
DB 1621 cattccggagaagatttttgcactacaagtttgatgaattccagattgcagatgttccat 1680
QY 1681 gatgcacaacatctccatctgagacccctcctaataatgagtcagagccgagcccaagctaaaccccc 1740
DB 1681 gatgcacaacatctccatctgagacccctcctaataatgagtcagagccgagcccaagctaaaccccc 1740
QY 1741 cgttccctctcgaacatacacactacatctccagataaataacatacactgacctcat 1800
DB 1741 cgttccctctcgaacatacacactacatctccagataaataacatacactgacctcat 1800
QY 1801 acttgcggagagacccaattaaagtgtgctgatttaacgatatctgcgaataccagatgagc 1860
DB 1801 acttgcggagagacccaattaaagtgtgctgatttaacgatatctgcgaataccagatgagc 1860
QY 1861 aacaacagtaacatacaactacatgcccagttccctgcagcagccgagagcagagctt 1920
DB 1861 aacaacagtaacatacaactacatgcccagttccctgcagcagccgagagcagagctt 1920
QY 1921 tgaagactcgttggagacacagttgctgcagctgcgcccagagacgcttactcagacaca 1980
DB 1921 tgaagactcgttggagacacagttgctgcagctgcgcccagagacgcttactcagacaca 1980
QY 1981 actgactgctcagtgctgcatacagaagaatgcttctatacaatgtgaaatgtgagac 2040
DB 1981 actgactgctcagtgctgcatacagaagaatgcttctatacaatgtgaaatgtgagac 2040
QY 2041 tttaaacctccctcctcctcctccttcaactcgcgaactgagttttcccatgtgttc 2100
DB 2041 tttaaacctccctcctcctcctccttcaactcgcgaactgagttttttcccatgtgttc 2100
QY 2101 cagacaaaggtgactataataatattagtgtttgcagagaaaaaa 2151

Db 2101 cagacaaagtgactataataattagtgcttgcagaaaaaaa 2151

RESULT 6
AAH90079
ID AAH90079 standard; CDNA: 3220 BP.
XX AAH90079;
XX
XX 01-OCT-2001 (first entry)
XX
XX Human bone marrow cDNA, SEQ ID NO: 323.
XX
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX antitumor; antibacterial; antifungal; anti-HIV; haemostatic;
XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; infection; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
XX Homo sapiens.
XX
XX MO200153453-A2.
XX
XX 26-JUL-2001.
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao Qa, Zhou P, Drmanac RT;
XX
XX WPI: 2001-488707/53.
XX P-PSDB: AAM00960.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 428; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
XX expressed in the bone marrow. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX infection, or may result from an autoimmune disorder, a coagulation
XX disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX suppression of an inflammatory response or treatment of a nervous
XX system disorder such as Alzheimer's disease. Detection of the presence
XX or increased expression of the polynucleotide or the protein it
XX encodes is useful for the diagnosis and/or prognosis of one
XX or more types of cancer. The polynucleotide and polypeptide can be
XX used as nutritional sources or supplements and in the screening of
XX chemical compounds as potential drugs.
XX
XX Sequence 3220 BP: 878 A; 725 C; 770 G; 847 T; 0 other;

Query Match 98.6%; Score 2120.8; DB 22; Length 3220;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

OY 1 tgcgttgctgctgctgcacaaaggcgcccgccgagagcgagccagtgcc 60

Db 59 tgcgttgctgctgctgcacaaaggcgcccgccgagagcgagccagtgcc 118

OY 61 tgcgagattatgacccgcgagcgctgctgcaagaagaagcactcagttatgaat 120

Db 119 tgcgagattatgacccgcgagcgctgctgcaagaagaagcactcagttatgaat 178

OY 121 tccctagaagagaagactgtaataatgacgaacccctcagaagataatacagaagaag 180

Db 179 tccctagaagagaagactgtaataatgacgaacccctcagaagataatacagaagaag 238

OY 181 aattcatttagaagacatacaacagctgacagactcgtcttaacccaagaagacagta 240

Db 239 aattcatttagaagacatacaacagctgacagactcgtcttaacccaagaagacagta 298

OY 241 tgtttaagaagcactgctatgaagaagcagtaaatgtgtgcccacaaacaaacttgccaat 300

Db 299 tgtttaagaagcactgctatgaagaagcagtaaatgtgtgcccacaaacaaacttgccaat 358

OY 301 ggcactccagatgattgtgcccagaacggaacactccagcaagctatgaaagaa 360

Db 359 ggcactccagatgattgtgcccagaacggaacactccagcaagctatgaaagaa 418

OY 361 aaggaactgtgtgcaataactttgaagcagtggtcagagtcagatcaagtgtgaattgtg 420

Db 419 aaggaactgtgtgcaataactttgaagcagtggtcagagtcagatcaagtgtgaattgtg 478

OY 421 gaacattctatcccaaatgtgtcattacccaacatgagcacaataactcgtatctaa 480

Db 479 gaacattctatcccaaatgtgtcattacccaacatgagcacaataactcgtatctaa 538

OY 481 cctatgttgcaagaagattcattacacgctctgcagactcgggagatgagtcatacgct 540

Db 539 cctatgttgcaagaagattcattacacgctctgcagactcgggagatgagtcatacgct 598

OY 541 gagaacattcgtcatcactgagatgccaactcactatgctcctggaactgtgtgcaag 600

Db 599 gagaacattcgtcatcactgagatgccaactcactatgctcctggaactgtgtgcaag 658

OY 601 gaatgtaacgagtgacactcgtatgacatgctgtgtggaagaagcttatcgagaagtgct 660

Db 659 gaatgtaacgagtgacactcgtatgacatgctgtgtggaagaagcttatcgagaagtgct 718

OY 661 aggaacagattcctgctgtgagagagcctgacgaagaagaagatgaggagacttatctc 720

Db 719 aggaacagattcctgctgtgagagagcctgacgaagaagaagatgaggagacttatctc 778

OY 721 aaaaaaacctctgacggaatgctcctcccaactcttttataagagacttatcct 780

Db 779 aaaaaaacctctgacggaatgctcctcccaactcttttataagagacttatcct 838

OY 781 aaaaattacaagaacattgagaacaatagaaatcatttggagatggaagacatagttta 840

Db 839 aaaaattacaagaacattgagaacaatagaaatcatttggagatggaagacatagttta 898

OY 841 cagagaattcctcctgacggaatggaacaaagaaagaaatttactggtttagagttatgat 900

Db 899 cagagaattcctcctgacggaatggaacaaagaaagaaatttactggtttagagttatgat 958

OY 901 cagagaattcctcctgacggaatggaacaaagaaagaaatttactggtttagagttatgat 960

Db 959 cagagaattcctcctgacggaatggaacaaagaaagaaatttactggtttagagttatgat 1018

OY 961 gaatgcaagcgaattctcacaagcgaacatagagttcagtcctgtctccagatgatgag 1020

Db 1019 gaatgcaagcgaattctcacaagcgaacatagagttcagtcctgtctccagatgatgag 1078

OY 1021 agagtgtacataaagaagatatacgattccacggttcagatgagtgagtgtaataacaggt 1080

Db 1079 agagtgtacataaagaagatatacgattccacggttcagatgagtgagtgtaataacaggt 1138

OY 1081 gaatgtcaaacacggttgatccacattgtgaagcaattcgcacactgcttcaaat 1140

```

Db 1139 gaaatgctaacaagcttgatccattgtgaagcaagcttcgacttcgtttcataat 1198
QY 1141 ggcatagtgtgactgctcccaaatcgcttcattgctgataagataagcctccca 1200
Db 1199 ggcatagtgtgactgctcccaaatcgcttcattgctgataagataagcctccca 1258
QY 1201 actgaattaccctcccgaggggtgctgctgcagcaaccgagcgtcgttaattgtaac 1260
Db 1259 actgaattaccctcccgaggggtgctgctgcagcaaccgagcgtcgttaattgtaac 1318
QY 1261 ttgtatgacaagatacattgttctgcattcgggataaactataaagatagacaaca 1320
Db 1319 ttgtatgacaagatacattgttctgcattcgggataaactataaagatagacaaca 1378
QY 1321 agtactgtgaattgttaagaccttaattgacacaaacgagcattgctgtttcag 1380
Db 1379 agtactgtgaattgttaagaccttaattgacacaaacgagcattgctgtttcag 1438
QY 1381 tacagggaagagctgtgtagtgcattcgcatactgacaacacatataatgtagaca 1440
Db 1439 tacagggaagagctgtgtagtgcattcgcatactgacaacacatataatgtagaca 1498
QY 1441 gaattgtgtgctgtttagcaggtcttaagaagcgaatggtgctgttattcga 1500
Db 1499 gaattgtgtgctgtttagcaggtcttaagaagcgaatggtgctgttattcga 1558
QY 1501 ttgtataacaagaagatagctagtgaggcctatgataagaaatgaagtgtgagatct 1560
Db 1559 ttgtataacaagaagatagctagtgaggcctatgataagaaatgaagtgtgagatct 1618
QY 1561 gtgctgcttggagcccgctgctcgcagggagacatctgtctcagacccttgtag 1620
Db 1619 gtgctgcttggagcccgctgctcgcagggagacatctgtctcagacccttgtag 1678
QY 1621 caticcggaagagcttttcgcatacagtttgaattcccgatgtcagatgttcaat 1680
Db 1679 caticcggaagagcttttcgcatacagtttgaattcccgatgtcagatgttcaat 1738
QY 1681 gatgacacaatcctcatctgtgacttccctaataatgatacagctgcgcaacgcccc 1740
Db 1739 gatgacacaatcctcatctgtgacttccctaataatgatacagctgcgcaacgcccc 1798
QY 1741 cgttcccttccgacatacaccatactccagataataaatacactgacatctat 1800
Db 1799 cgttcccttccgacatacaccatactccagataataaatacactgacatctat 1858
QY 1801 actgcccagagaccattaaag-ttgcglatiaacgtatctcccaatacagagatag 1859
Db 1859 actgcccagagaccattaaag-ttgcglatiaacgtatctcccaatacagagatag 1918
QY 1860 caacaacagtaacaatacactactgcccagttccctgtagcagcgaagagagagct 1919
Db 1919 caacaacagtaacaatacactactgcccagttccctgtagcagcgaagagagagct 1978
QY 1920 ttgaagcctcgttggagcacagttgtgtcagttcgcccgagacaggtctaccacagac 1979
Db 1979 ttgaagcctcgttggagcacagttgtgtcagttcgcccgagacaggtctaccacagac 2038
QY 1980 aactgactgtcagttgtctatcaagaagaagcttcttacaatgtgaattgaa 2039
Db 2039 aactgactgtcagttgtctatcaagaagaagcttcttacaatgtgaattgaa 2098
QY 2040 ctttaaacctccctcctcctccttccaaactctgaactagtttttccatggt 2099
Db 2099 ctttaaacctccctcctcctccttccaaactctgaactagtttttccatggt 2158
QY 2100 ccagacaagaagtgaactataataatatttagtgtttgccaaga 2143
Db 2159 ccagacaagaagtgaactataataatatttagtgtttgccaaga 2202

```

RESULT 7
AAZ29233

```

ID AAZ29233 standard; cDNA; 2419 BP.
XX
AC AAZ29233;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human cell signalling protein-12 encoding cDNA.
XX
KW Cell signalling protein-12; CSIGP-12; cell proliferation;
KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
KW arteriosclerosis; Addison's disease; multiple sclerosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 70..1779
FT /tag= a
FT /product= "Cell Signalling Protein-12"
XX
PN WO958558-A2.
XX
PD 18-NOV-1999.
XX
PF 13-MAY-1999; 99MO-US10567.
XX
PR 13-MAY-1998; 98US-0085343.
XX
PR 26-AUG-1998; 98US-0098010.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
PI Baughn MR, Yang J;
XX
DR WPI: 2000-086432/07.
DR P-PSDB: AAY44249.
XX
PT Human cell signaling proteins useful for, e.g. diagnosing cell
PT proliferative and inflammatory disorders
XX
PS Claim 9; Page 87-88; 90pp; English.
XX
CC The present sequence is a cDNA obtained from Incyte clone 3239149 of
CC COLACTO1 library. It encodes cell signalling protein-12 (CSIGP-12). It
CC is expressed in musculo-skeletal, gastrointestinal and nervous
CC tissues. Fragments of CSIGP encoding nucleic acid can be used as
CC hybridisation probe for detecting CSIGP related sequences or allelic
CC variants. Recombinant CSIGP can be produced in host cells by transforming
CC them with genetically engineered vectors. Agonists or antagonists can be
CC used in the treatment of cell proliferative and inflammatory disorders
CC associated with decreased or increased CSIGP expression. CSIGP is used in
CC the diagnosis, prevention or treatment of cell proliferative disorders
CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.
XX
SQ Sequence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other:

```

Query Match 98.3%; Score 2114.8; DB 21; Length 2419;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2119; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 11 gcgagctgacccaaggcgcccgcgagagcgagaccagtgctcgcgatga 70
Db 11 ggggctgacccaaggcgcccgcgagagcgagaccagtgctcgcgatga 70
QY 71 tggaccgagcgagggcggtgctgcaagaagaagcactcaagttatgaattcctaaga 130
Db 71 tggaccgagcgagggcggtgctgcaagaagaagcactcaagttatgaattcctaaga 130
QY 131 gagaagactgataatgtgcgaaccctcaggaagataatacagaagaattacta 190
Db 131 gagaagactgataatgtgcgaaccctcaggaagataatacagaagaattacta 190

```

QY 191 gacagacatacaacagctgtgccaagcttgcttaacaaagaacagatgtttagca 250
|||||
Db 191 gacagacatacaacagctgtgccaagcttgcttaacaaagaacagatgtttagca 250
QY 251 gacactctatgaagactgaatgtgtgtgccaacaaactgtgcaatgtgcaatcca 310
|||||
Db 251 gacactctatgaagactgaatgtgtgtgccaacaaactgtgcaatgtgcaatcca 310
QY 311 gtaatgtgtgtgccaacaaagaaactctcagcaagctatgaaagaaagaaactgt 370
|||||
Db 311 gtaatgtgtgtgccaacaaagaaactctcagcaagctatgaaagaaagaaactgt 370
QY 371 ggtgcaataactgtgcaagctgtgcaagctgaatcagatcagatgtgtgtggaactctta 430
|||||
Db 371 ggtgcaataactgtgcaagctgtgcaagctgaatcagatcagatgtgtgtggaactctta 430
QY 431 tatcccaaatgtgtcaatacacaatgtggaacataaactcgtatcttaacctatgtgtc 490
|||||
Db 431 tatcccaaatgtgtcaatacacaatgtggaacataaactcgtatcttaacctatgtgtc 490
QY 491 agaggaattcataactgtgtgtgccaagctcggggaattgtatcatatcgtctggaacatcc 550
|||||
Db 491 agaggaattcataactgtgtgtgccaagctcggggaattgtatcatatcgtctggaacatcc 550
QY 551 tgcataactgtgatagtgccaatactactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 610
|||||
Db 551 tgcataactgtgatagtgccaatactactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 610
QY 611 gaatgacctctgatagtgcatgtgtgtgaaagaacttaacagagaatgtgtgtgtgtgtgtgt 670
|||||
Db 611 gaatgacctctgatagtgcatgtgtgtgaaagaacttaacagagaatgtgtgtgtgtgtgtgt 670
QY 671 ctctgtgtggaagctgt 730
|||||
Db 671 ctctgtgtggaagctgt 730
QY 731 ctctgtgtggaagctgt 790
|||||
Db 731 ctctgtgtggaagctgt 790
QY 791 aagacatactgagaacatagatcttaattgtgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 850
|||||
Db 791 aagacatactgagaacatagatcttaattgtgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 850
QY 851 actgt 910
|||||
Db 851 actgt 910
QY 911 taagcgacctgtgagaac 970
|||||
Db 911 taagcgacctgtgagaac 970
QY 971 gaatttcacagagcaataaagttcagtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1030
|||||
Db 971 gaatttcacagagcaataaagttcagtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1030
QY 1031 taacagatcatcagatctccacaggttcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1090
|||||
Db 1031 taacagatcatcagatctccacaggttcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1090
QY 1091 aacaggt 1150
|||||
Db 1091 aacaggt 1150
QY 1151 tgacctgt 1210
|||||
Db 1151 tgacctgt 1210
QY 1211 cccctcggagaggt 1270
|||||
Db 1211 cccctcggagaggt 1270
QY 1271 agtaacatgttctgtcatctgt 1330
|||||

Db 1271 agtaacatgttctgtcatctgt 1330
QY 1331 aattgttaaggaaccttaattgt 1390
|||||
Db 1331 aattgttaaggaaccttaattgt 1390
QY 1391 ggt 1450
|||||
Db 1391 ggt 1450
QY 1451 catgt 1510
|||||
Db 1451 catgt 1510
QY 1511 agaggaatgt 1570
|||||
Db 1511 agaggaatgt 1570
QY 1571 tgagccccgt 1630
|||||
Db 1571 tgagccccgt 1630
QY 1631 gatttttgcagtaacgt 1690
|||||
Db 1631 gatttttgcagtaacgt 1690
QY 1691 tccatcatctgt 1750
|||||
Db 1691 tccatcatctgt 1750
QY 1751 ctgtgaacatacactgt 1810
|||||
Db 1751 ctgtgaacatacactgt 1810
QY 1811 gaccttaaaagttgt 1870
|||||
Db 1811 gaccttaaaagttgt 1870
QY 1871 acaatcaactactgt 1930
|||||
Db 1871 acaatcaactactgt 1930
QY 1931 gttgt 1990
|||||
Db 1931 gttgt 1990
QY 1991 tcaagt 2050
|||||
Db 1991 tcaagt 2050
QY 2051 cccctcct 2110
|||||
Db 2051 cccctcct 2110
QY 2111 tgacttaataatataatgt 2136
|||||
Db 2111 tgacttaataatataatgt 2136

RESULT 8
AAH89966
ID AAH89966 standard; cDNA; 3622 BP.
XX
XX AAH89966;
AC
XX
XX 01-Oct-2001 (first entry)
DT
XX
XX Human bone marrow cDNA, SEQ ID NO: 97.
DE
DE Human; bone marrow; antiinflammatory; cytosolic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;

| | | | |
|----------|---|---|------|
| Db | 1581 | tcggagacaagaatgtggtgtcgtttacagaggtgttgaagaagccatgaggaattcgtgcgt | 1640 |
| Qy | 1492 | tgatcttgatttgataacaagaaggaatgcaatgagggccatgatatggaataaataaagtg | 1551 |
| Db | 1641 | tgatctcgatttgataacaagaaggaatgacagtgaggccattatggaataaataaagtg | 1700 |
| Qy | 1552 | tggaatccttgaggtgtgtcttgagcccccggtgcctcctcagaagaacatgctctcagacc | 1611 |
| Db | 1701 | tggaatccttgaggtgtgtcttgagcccccggtgcctcctcagaagaacatgctctcagacc | 1760 |
| Qy | 1612 | ctgtgtgagacatcccggaagaagtttttcgactacagtttgatgtaattccagattgtcagt | 1671 |
| Db | 1761 | ctgtgtgagacatcccggaagaagtttttcgactacagtttgatgtaattccagattgtcagt | 1820 |
| Qy | 1672 | agttcacatgtatgacacaatccctatcttgagacttccctaataatgatacagctgccagct | 1731 |
| Db | 1821 | agttcacatgtatgacacaatccctatcttgagacttccctaataatgatacagctgccagct | 1880 |
| Qy | 1732 | gaacccccccgttccctctctcgacacatacactacatctccagataaataacatacac | 1791 |
| Db | 1881 | gaacccccccgttccctctctcgacacatacactacatctccagataaataacatacac | 1940 |
| Qy | 1792 | tgacctcaactgtgccagagaccatataag-ttgcggtattaaagtgatctgcgaatac | 1850 |
| Db | 1941 | tgacctcaactgtgccagagaccatataagtttgcggtattaaagtgatctgcgaatac | 2000 |
| Qy | 1851 | caggaatgacacacaacagttaacaatacaactgcgccagtttccctggagctagccaggt | 1910 |
| Db | 2001 | caggaatgacacacaacagttaacaatacaactgcgccagtttccctggagctagccaggt | 2060 |
| Qy | 1911 | agcagaggtcttgagaactcgtgttggagacaagttgtgtctgcagtcggccaggaagctc | 1970 |
| Db | 2061 | agcagaggtcttgagaactcgtgttggagacaagttgtgtctgcagtcggccaggaagctc | 2120 |
| Qy | 1971 | actcagacacaactgactcctcctcagtgctgcgtacacagaagaatgctcttatcaatttgaa | 2030 |
| Db | 2121 | actcagacacaactgactcctcctcagtgctgcgtacacagaagaatgctcttatcaatttgaa | 2180 |
| Qy | 2031 | tgattggaacttttaaacctccctcctcctcctcctccttaccctctgcacactagtttttc | 2090 |
| Db | 2181 | tgattggaacttttaaacctccctcctcctcctcctcctccttaccctctgcacactagtttttc | 2240 |
| Qy | 2091 | ccattggtgtccagacaaaagtgacttaataatataattagtggtttccagaa | 2143 |
| Db | 2241 | ccattggtgtccagacaaaagtgacttaataatataattagtggtttccagaa | 2293 |
| RESULT 9 | | | |
| AAK52699 | | | |
| ID | AAK52699 standard; CDNA: 3003 BP. | | |
| XX | AAK52699; | | |
| AC | | | |
| XX | 06-NOV-2001 (first entry) | | |
| DT | | | |
| XX | Human polynucleotide SEQ ID NO 2228. | | |
| DE | | | |
| XX | | | |
| KM | Human; cytokine; cell proliferation; cell differentiation; gene therapy | | |
| KM | vaccine; peptide therapy; stem cell growth factor; haematopoiesis; | | |
| KW | tissue growth factor; immunomodulatory; cancer; leukaemia; | | |
| KM | nervous system disorder; arthritis; inflammation; ss. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PM | WO200157190-A2. | | |
| XX | | | |
| PD | 09-AUG-2001. | | |
| XX | | | |
| PF | 05-FEB-2001; 2001WO-US04098. | | |
| XX | | | |
| PR | 03-FEB-2000; 2000US-0496914. | | |
| PR | 27-APR-2000; 2000US-0360875. | | |
| PR | 20-JUN-2000; 2000US-03598075. | | |

| Query Match | 93.4% | Score 2008, 2: | DB 22: | Length 3003: |
|-----------------------|--------------|--|---------------|-------------------|
| Best Local Similarity | 99.6% | Pred. No. 0: | | |
| Matches 2024: | Conservative | 0: | Mismatches 8: | Indels 1: Gaps 1. |
| QY | 112 | ttatgaatcttcctgagagagagaagacgttaataatgagcccccctagagaataata | 171 | |
| DB | 298 | ttccagaatctccctcagagagagaagcgttaataatgagcccccctagagaataata | 357 | |
| QY | 172 | ccagagaagatctcactcttgacagacaataacaacagctgycagagactcgtcttaacc | 231 | |
| DB | 358 | ccagagaagatctcactcttgacagacaataacaacagctgycagagactcgtcttaacc | 417 | |
| QY | 232 | gaacacgtatgttttagcaagacactgtctataagaagctggaattgtgtgccaataa | 291 | |
| DB | 418 | gaacacgtatgttttagcaagacactgtctataagaagctggaattgtgtgccaataa | 477 | |
| QY | 292 | cttgcacaatgagcaacttcagatagatgttggcccaagaacgggaactcccaagaag | 351 | |
| DB | 478 | cttgcacaatgagcaacttcagatagatgttggcccaagaacgggaactcccaagaag | 537 | |
| QY | 352 | gaaaagaaaaggaactgtgtgtccaataactcttgagcagtggttcagatccagatcag | 411 | |
| DB | 538 | gaaaagaaaaggaactgtgtgtccaataactcttgagcagtggttcagatccagatcag | 597 | |
| QY | 412 | gaattgtggaacatcttatccccaatggtgtcattaccacaacatggtgcacataaact | 471 | |
| DB | 598 | gaattgtggaacatcttatccccaatggtgtcattaccacaacatggtgcacataaact | 657 | |
| QY | 472 | tatcttaaacctatgtttgacagagagattcaataactgctctgcagctccggggattg | 531 | |
| DB | 658 | tatcttaaacctatgtttgacagagagattcaataactgctctgcagctccggggattg | 717 | |
| QY | 532 | catatgcctgagaaactttctgataccctggatggccaataactatagtctgtgaact | 591 | |
| DB | 718 | catatgcctgagaaactttctgataccctggatggccaataactatagtctgtgaact | 777 | |

| CC | activin/inhibin activity and may be useful in the diagnosis and/or |
|----------------------------|---|
| CC | treatment of cancer, leukaemia, nervous system disorders, arthritis and |
| CC | inflammation. |
| CC | Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 |
| CC | (AAM80020) are omitted as the relevant pages from the sequence listing |
| CC | were missing at the time of publication. |
| XX | |
| SO | Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other; |
| Query Match | 93.4%; Score 2008.2; DB 22; Length 3003; |
| Best Local Similarity | 99.6%; Pred. No. 0; |
| Matches 2024; Conservative | 0; Mismatches 8; Indels 1; Gaps 1. |
| QY | 112 ttctgaattccctcagagagagagactgttaataatgtagaccccttagaagaatata 171 |
| | |
| | |
| Db | 298 ttccagattccctcagagagagagactgttaataatgtagaccccttagaagaatata 357 |
| QY | 172 ccagagaagaattcacttagacagacatacaacagctgtgcagacactgtcttaaacca 231 |
| | |
| | |
| Db | 358 ccagagaagaattcacttagacagacatacaacagctgtgcagacactgtcttaaacca 417 |

| | | | |
|----|------|--|------|
| OY | 112 | ccaaagaaagattccacttgcagaaatatacaacacgtctgccaactctcttaaacca | 231 |
| Db | 358 | ccagagaagattccactagacagacatacaacacgctctgcagactctgttaaacca | 417 |
| OY | 232 | gaacacgtatggttaagcaagcactgtctatgaagactggaattgtgtgccaaaacaaa | 291 |
| Db | 418 | gaacaactatgttttagcaagcactgtctatgaagacttgaattgtgtgccaaaacaaa | 477 |
| OY | 292 | cttcccaatggaactccagatgatgtgtgtgcccaagcaaggaaccttaagcaactat | 351 |
| Db | 478 | cttgcacatggaactccagatgatgtgtgtgcccaagcaaggaaccttaagcaactat | 537 |
| OY | 352 | gaaaggaagaaagaaactgtctgtlcaaatactttgacagtggtlcaagatcagatacgt | 411 |
| Db | 538 | gaaaggaagaaagaaactgtctgtlcaaatactttgacagtggtlcaagatcagatacgt | 597 |
| OY | 412 | gaattgttgaaactcttatatcccacaattgttcaatccacaactgtggcacataactcg | 471 |
| Db | 598 | gaattgttgaaactcttatatcccacaattgttcaatccacaactgtggcacataactcg | 657 |
| OY | 472 | tattttaaactatgtctgcagagagatttcaataactgtctgcagactcgggattggat | 531 |
| Db | 658 | tattttaaactatgtctgcagagagatttcaataactgtctgcagactcgggattggat | 717 |
| OY | 532 | catatcgctgaagaacattctgtcalactgtgagtcgaactaactatgtctgttgaact | 591 |
| Db | 718 | catatgtctgaagaacattctgtcalactgtgagtcgaactaactatgtctgttgaact | 777 |
| OY | 592 | gtgtgcaaggaatggtgacccgagtgacccctctgatgtgacgtctgtgaaagaactatcag | 651 |
| Db | 778 | gtgtgcaaggaatggtgacccgagtgacccctctgatgtgacgtctgtgaaagaactatcag | 837 |
| OY | 652 | agaatgttcaggaacattctctgtgagagagccctgtgagagacgaagaagatgtggacag | 711 |
| Db | 838 | agaatgttcaggaacattctctgtgagagagccctgtgagagacgaagaagatgtggacag | 897 |
| OY | 712 | tatttatcaaaaaaacaacctctctgcagtgagtgctccctcccaactcttttataagaa | 771 |
| Db | 898 | tatttatcaaaaaaacaacctctctgcagtgagtgctccctcccaactcttttataagaa | 957 |
| OY | 772 | ctttatccctaataattatacaagaactttgagacatatgaattctaatgtgagatvtggaaga | 831 |
| Db | 958 | ctttatccctaataattatacaagaactttgagacatatgaattctaatgtgagatvtggaaga | 1017 |
| OY | 832 | catagtttcaaggaattcactgcgcgaagtgaacaacagcaaggaacttactctttacag | 891 |
| Db | 1018 | catagtttcaaggaattcactgcgcgaagtgaacaacagcaaggaacttactctttacag | 1077 |
| OY | 892 | tatgatgatcagaanaatagtaagcggcctctlogagacacaacatcaagatcttvgatataa | 951 |
| Db | 1078 | tatgatgatcagaanaatagtaagcggcctctlogagacacaacatcaagatcttvgatataa | 1137 |
| OY | 952 | aaacaatttgaaatgtcagaagcgaattctcaagcgcatcacaggttcaagtcctctgtctcag | 1011 |
| Db | 1138 | aaacaatttgaaatgtcagaagcgaattctcaagcgcatcacaggttcaagtcctctgtctcag | 1197 |

```

QY 1012 tatgatgagagatgatacataacagatcatcgattccacggtcgagtgctgagatga 1071
    |||||
Db 1198 tatgatgagagagatgatacataacagatcatcgattccacggtcgagtgctgagatga 1257
QY 1072 aatacagtgaaatgataacacgattgattcaccattgtgaagcagttctgcactgcgt 1131
    |||||
Db 1258 aatacagtgaaatgataacacgattgattcaccattgtgaagcagttctgcactgcgt 1317
QY 1132 ttaataaagagcagtgatggtgacgttcacaaagatcggttcacgtgtgtatggagatga 1191
    |||||
Db 1318 ttaataaagagcagtgatggtgacgttcacaaagatcggttcacgtgtgtatggagatga 1377
QY 1192 gctccccaactacattacccctccgagagtgctgtgtcgacacccgagctgctgcaat 1251
    |||||
Db 1378 gctccccaactacattacccctccgagagtgctgtgtcgacacccgagctgctgcaat 1437
QY 1252 gttgtgacttgatgataacagatcatgtttctgcactctggggatagaaactataagata 1311
    |||||
Db 1438 gttgtgacttgatgataacagatcatgtttctgcactctggggatagaaactataagata 1497
QY 1312 tggacaacagatcctgtgattgtgaagacctaataatggacacaaacgagcattgccc 1371
    |||||
Db 1498 tggacaacagatcctgtgattgtgaagacctaataatggacacaaacgagcattgccc 1557
QY 1372 tgtttgagtagacagagacagagctgtgagtgaagtgtgctcattgacacacatacagatta 1431
    |||||
Db 1558 tgtttgagtagacagagacagagctgtgagtgaagtgtgctcattgacacacatacagatta 1617
QY 1432 tggagacatgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1491
    |||||
Db 1618 tggagacatgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1677
QY 1492 tgcattcgattgtataacaaagagatgatacagtgaggggccatataatgaaataatgaatga 1551
    |||||
Db 1678 tgcattcgattgtataacaaagagatgatacagtgaggggccatataatgaaataatgaatga 1737
QY 1552 tgggattctgtgctgctgtgtgaaccccgctgtcctgtcagaggaacactctgtctagagac 1611
    |||||
Db 1738 tgggattctgtgctgctgtgtgaaccccgctgtcctgtcagaggaacactctgtctagagac 1797
QY 1612 ctgtgtgagatcccggaagaagttttcgactcagtttgatgaattccagattgtcagat 1671
    |||||
Db 1798 ctgtgtgagatcccggaagaagttttcgactcagtttgatgaattccagattgtcagat 1857
QY 1672 agttcacatgatgacaaacacccatcctcgtgagactcctcctaataatcagctgagccaaagt 1731
    |||||
Db 1858 agttcacatgatgacaaacacccatcctcgtgagactcctcctaataatcagctgagccaaagt 1917
QY 1732 gaaccccccgttcctcctctcgaacatacactacatcctcgaataaataacatacac 1791
    |||||
Db 1918 gaaccccccgttcctcctctcgaacatacactacatcctcgaataaataacatacac 1977
QY 1792 tgacctatattctggccagagacccattaaag-cttgcggtatttaacgtattctccatatac 1850
    |||||
Db 1978 tgacctatattctggccagagacccattaaag-cttgcggtatttaacgtattctccatatac 2037
QY 1851 cagatgtagacaacacagataacatacactactgcccagtttccctggagctagccgagag 1910
    |||||
Db 2038 cagatgtagacaacacagataacatacactactgcccagtttccctggagctagccgagag 2097
QY 1911 agcagagcttgagactcctgtttgggacacagttgctcagtcagtcggccagagagctct 1970
    |||||
Db 2098 agcagagcttgagactcctgtttgggacacagttgctcagtcagtcggccagagagctct 2157
QY 1971 actaagacacactgagctgttcagttgtgtatcaagaagtgtcttcatcaattgtgaa 2030
    |||||
Db 2158 actaagacacactgagctgttcagttgtgtatcaagaagtgtcttcatcaattgtgaa 2217
QY 2031 tgattgaaactttaaacctccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2090
    |||||
Db 2218 tgattgaaactttaaacctccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2277

```

```

QY 2091 coattggtctcagacaaagtgacttataatataattagttgttctgcagaa 2143
    |||||
Db 2278 coattggtctcagacaaagtgacttataatataattagttgttctgcagaa 2330

```

RESULT 12

AAA73132
ID AAA73132 standard; cDNA: 1707 BP.

AC AAA73132;

DT 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) cDNA.

KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;

KW beta-catenin; SKP1; Cull1; F-box motif; WD40 repeat motif; FMD1;

KW gene therapy; colon cancer; beta-transducin repeat containing protein;

KW beta-TrCP; ss.

OS Homo sapiens.

PN JP200016542-A.

PD 20-JUN-2000.

PF 02-DEC-1998; 98JP-0343437.

PR 02-DEC-1998; 98JP-0343437.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

DR WPI, 2000-485550/43.

DR P-PSDB; AAB12813.

PT F-box protein of ubiquitin ligase SCF complex which promotes the

PT ubiquitination of IkappaB or beta-catenin

PS Disclosure: Fig 17; 19pp; Japanese.

CC The present invention describes an F-box motif protein of ubiquitin

CC ligase SCF complex which promotes the ubiquitination of IkappaB or

CC beta-catenin and is constituted by SKP1 protein, Cull1 protein and a

CC complex (SCF complex) of F-box protein containing F-box motif and WD40

CC repeat motif and has the amino acid sequence of 43 residues (AAB12813)

CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin

CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin

CC repeat containing protein (beta-TrCP)). The F-box protein can be used for

CC the gene therapy of colon cancer by being recombined to a virus vector.

CC The present sequence encodes the human beta-TrCP protein from the

CC present invention.

CC

SO Sequence 1707 BP; 514 A; 349 C; 414 G; 430 T; 0 other;

Query Match 79.4%; Score 1707; DB 21; Length 1707;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 70 atggaccggcgagcggtgctgcaaggaagcaccacgaattatgattcctcagag 129
    |||||
Db 1 atggaccggcgagcggtgctgcaaggaagcaccacgaattatgattcctcagag 60
QY 130 agagaagactgtataatgycgaaccccttagaagataatcagagaagaattcact 189
    |||||
Db 61 agagaagactgtataatgycgaaccccttagaagataatcagagaagaattcact 120
QY 190 agacagacatacaacagctgtgcccagactcgtcttaaaccaagaacagatgattgaa 249
    |||||
Db 121 agacagacatacaacagctgtgcccagactcgtcttaaaccaagaacagatgattgaa 180
QY 250 agcactgctataaagactagaaattgtgtgccaacaaacactgccaatggcactcc 309
    |||||

```

Db 181 agcactgctatgaagactggaattgtgtgcccacaaacaaacttgccaatgagcacttcc 240
 QY 310 agtatgatttgcccaagcagcaagaaactctcagcaagctatgaataaagaaagaaactg 369
 Db 241 agtatgatttgcccaagcagcaagaaactctcagcaagctatgaataaagaaagaaactg 300
 QY 370 tgtgtcaaatctctgagcagtgctcagagtcagatcaagtggaatttgtagaacctct 429
 Db 301 tgtgtcaaatctctgagcagtgctcagagtcagatcaagtggaatttgtagaacctct 360
 QY 430 atatacccaaatgtgtcattaccacaatgagcacataactcgtatcttaacctatgtgtg 489
 Db 361 atatcccaaatgtgtcattaccacaatgagcacataactcgtatcttaacctatgtgtg 420
 QY 490 caaagagatttcaactgtctgcagcctcgaggttgatgataatgcctgagaaact 549
 Db 421 caaagagatttcaactgtctgcagcctcgaggttgatgataatgcctgagaaact 480
 QY 550 ctgtcataccttgatcccaaatcaactatgtgtctgtgaacttggtgtcaagaaatgtgtac 609
 Db 481 ctgtcataccttgatcccaaatcaactatgtgtctgtgaacttggtgtcaagaaatgtgtac 540
 QY 610 cgaagtacactctgatggcctgtgtgtgaagaagcttatacgaagaatgtgtcaagaat 669
 Db 541 cgaagtacactctgatggcctgtgtgtgaagaagcttatacgaagaatgtgtcaagaat 600
 QY 670 tctctgttgagagagcctgtgcagacgaagaagatgagtgagcaatattatccaacaa 729
 Db 601 tctctgttgagagagcctgtgcagacgaagaagatgagtgagcaatattatccaacaa 660
 QY 730 cctcctcgaggggaatgtcctcccaactctttatagaagcacttatacctaataata 789
 Db 661 cctcctcgaggggaatgtcctcccaactctttatagaagcacttatacctaataata 720
 QY 790 caagacatgagacaaatagaatctaatctgagatgtgagaaacatagtttacagaagaatt 849
 Db 721 caagacatgagacaaatagaatctaatctgagatgtgagaaacatagtttacagaagaatt 780
 QY 850 cactgcggaagtgaacaaagcaagaggttactcgtttacaglatgatlacagaanaata 909
 Db 781 cactgcggaagtgaacaaagcaagaggttactcgtttacaglatgatlacagaanaata 840
 QY 910 gtaagagccttcggaacacacaaatcaagatctgggataaaacacatctggaatgcaag 969
 Db 841 gtaagagccttcggaacacacaaatcaagatctgggataaaacacatctggaatgcaag 900
 QY 970 cgaattctacagagccatacaggttcagtcctctgtctccatgatacgaagaagatgac 1029
 Db 901 cgaattctacagagccatacaggttcagtcctctgtctccatgatacgaagaagatgac 960
 QY 1030 ataacagagatcatcgattccacaggtcagagtggtgagatgataatacagtgaaatgtcta 1089
 Db 961 ataacagagatcatcgattccacaggttcagagtggtgagatgataatacagtgaaatgtcta 1020
 QY 1090 aacacgttgatcacacacattgtgaagcaggttcgacacttgcttttaataatgagcagtg 1149
 Db 1021 aacacgttgatcacacacattgtgaagcaggttcgacacttgcttttaataatgagcagtg 1080
 QY 1150 gtgacctgtctccaaagatcgttccatctgtatagatgataatgacctcccaactgaact 1209
 Db 1081 gtgacctgtctccaaagatcgttccatctgtatagatgataatgacctcccaactgaact 1140
 QY 1210 accctcggaaggtgtgtgtgagacacgaggtgtgttccaatgtgtgagacttggatgac 1269
 Db 1141 accctcggaaggtgtgtgtgagacacgaggtgtgttccaatgtgtgagacttggatgac 1200
 QY 1270 aagtaacattgttctgcatctgagatgagaactataaagatgataagacacaaactgtgt 1329
 Db 1201 aagtaacattgttctgcatctgagatgagaactataaagatgataagacacaaactgtgt 1260
 QY 1330 gaatttgaagacacataatgagacacaaagagcattgtcctgttttgcagtaacagagac 1389
 Db 1261 gaatttgaagacacataatgagacacaaagagcattgtcctgttttgcagtaacagagac 1320

QY 1390 aggcctggtgaatgtgtcctatctgacaaactatcagatataatgagacataagaatgtgt 1449
 Db 1321 aggcctggtgaatgtgtcctatctgacaaactatcagatataatgagacataagaatgtgt 1380
 QY 1450 gcatgttaacagatgttgaagagccatgaggaatttggtgtgttatcgaatttgataac 1509
 Db 1381 gcatgtttacagagtggttgaagagccatgaggaatttggtgtgttatcgaatttgataac 1440
 QY 1510 aagaggaatgacagtgaggccatgatatgaaataaataaagatgtgtggaatctgtgtgtct 1569
 Db 1441 aagaggaatgacagtgaggccatgatatgaaataaataaagatgtgtggaatctgtgtgtct 1500
 QY 1570 ttggaaccccggtgtcctcgcagaggaacactcgttataagagccctgtgtgagcattccgga 1629
 Db 1501 ttggaaccccggtgtcctcgcagaggaacactcgttataagagccctgtgtgagcattccgga 1560
 QY 1630 agagttttcgaactacagtttgatgaattccagattgtcagtagttcacatgatacaca 1689
 Db 1561 agagttttcgaactacagtttgatgaattccagattgtcagtagttcacatgatacaca 1620
 QY 1690 atctcatctgaggacttccataatgatccagctgtgcccagctgaaccccggttccct 1749
 Db 1621 atctcatctgaggacttccataatgatccagctgtgcccagctgaaccccggttccct 1680
 QY 1750 tctcgaacatcacctacatctccaga 1776
 Db 1681 tctcgaacatcacctacatctccaga 1707

RESULT 13
 AAK51715
 ID AAK51715 standard; cDNA; 2285 BP.
 XX
 AC AAK51715;
 XX
 DT 06-NOV-2001 (first entry)
 DE
 DE Human polynucleotide SEQ ID NO 260.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 PD
 PD 09-AUG-2001.
 PF
 PF 05-FEB-2001; 2001WO-US04098.
 PR
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejman T, Goodrich R;
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM78582.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -

AC AAK51717;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 262.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001: 2001WO-US04098.
XX
XX 03-FEB-2000: 2000US-0496914.
PR 27-APR-2000: 2000US-0560875.
PR 20-JUN-2000: 2000US-0598075.
PR 19-JUL-2000: 2000US-0620325.
PR 01-SEP-2000: 2000US-0654936.
PR 15-SEP-2000: 2000US-0663561.
PR 20-OCT-2000: 2000US-0693325.
PR 30-NOV-2000: 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dermanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
DR P-PSDB: AAM78584.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1: Page 1183-1186; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
SQ Sequence 2366 BP; 661 A; 519 C; 594 G; 592 T; 0 other;

Query Match 78.9%: Score 1697.8; DB 22: Length 2366;
Best Local Similarity 99.0%: Pred. No. 0;
Matches 1708: Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 112 ttatgaattcctcagagagaagactgtataatgcaaccccttagaagatatata 171
Db 358 ttcatgaattcctcagagagaagactgtataatgcaaccccttagaagatatata 417
QY 172 ccagagaagaattcacttagacacatacaacagctgtgcccagactgtcttaaacaa 231
Db 418 ccagagaagaattcacttagacacatacaacagctgtgcccagactgtcttaaacaa 477
QY 232 gaacagatgttttagcaagacacgctcatgaagactgagaattgtgtgcccacaa 291

Db 478 gaacagatgttttagcaagacacgctcatgaagactgagaattgtgtgcccacaa 537
QY 292 ctgccaatggaacttcacagatgatgtgtgcccagcaagaaactctcagaagctat 351
Db 538 ctgccaatggaacttcacagatgatgtgtgcccagcaagaaactctcagaagctat 597
QY 352 gaagaagaaaaggaactgtgtgtcaaatcacttgaagcagtggtcagaagtcagaagta 411
Db 598 gaagaagaaaaggaactgtgtgtcaaatcacttgaagcagtggtcagaagtcagaagta 657
QY 412 gaatttggacacatctatatacccaatgtgtcattacccaatgtgtgccaataactcg 471
Db 658 gaatttggacacatctatatacccaatgtgtcattacccaatgtgtgccaataactcg 717
QY 472 tatcttaacctatgtgtcagagagattcataactgtcttcagactcgggattgagat 531
Db 718 tatcttaacctatgtgtcagagagattcataactgtcttcagactcgggattgagat 777
QY 532 catatgcctgagaacattcctgtatatactgtgatgccaataactatgtctgtgaact 591
Db 778 catatgcctgagaacattcctgtatatactgtgatgccaataactatgtctgtgaact 837
QY 592 gtgtgcaaggaatgtgtacccagatgacccctcgtatggtcagctgtgtggaagaagctatcag 651
Db 838 gtgtgcaaggaatgtgtacccagatgacccctcgtatggtcagctgtgtggaagaagctatcag 897
QY 652 agaatgtgtcagagacattctctgtgtgagagccctgtgcaagcaagagatgtggagacag 711
Db 898 agaatgtgtcagagacattctctgtgtgagagccctgtgcaagcaagagatgtggagacag 957
QY 712 tattattcaaaaaaacactcctcgtacgggaatggtccctcccaactctttataagaga 771
Db 958 tattattcaaaaaaacactcctcgtacgggaatggtccctcccaactctttataagaga 1017
QY 772 cttatccctaaatltacaaagacattgagacaaataagaaatcgaatgtgagatgtggaaga 831
Db 1018 cttatccctaaatltacaaagacattgagacaaataagaaatcgaatgtgagatgtggaaga 1077
QY 832 catagatttacagagaattcactcgcgaagtgtgaaacagaagagattacgtttacag 891
Db 1078 catagatttacagagaattcactcgcgaagtgtgaaacagaagagattacgtttacag 1137
QY 892 tatgatgatcagaanaatagtaagcggccttcgagaaacaacaatacagaatctgtgataaa 951
Db 1138 tatgatgatcagaanaatagtaagcggccttcgagaaacaacaatacagaatctgtgataaa 1197
QY 952 aacacattggaatgtcagaagcaatctcgaagcgaatacagaagttcagctcgtgtccag 1011
Db 1198 aacacattggaatgtcagaagcaatctcgaagcgaatacagaagttcagctcgtgtccag 1257
QY 1012 tatgagaagagatgtatcatcaacagatcatcgattccacgggttcagatgtgtgagatga 1071
Db 1258 tatgagaagagatgtatcatcaacagatcatcgattccacgggttcagatgtgtgagatga 1317
QY 1072 aatcacaggtgaatgtcacaacagttgatccacattgtgaagcagtttcgaactgtcgt 1131
Db 1318 aatcacaggtgaatgtcacaacagttgatccacattgtgaagcagtttcgaactgtcgt 1377
QY 1132 ttcaataatgtgcatgtgtgtgacgtgctccaaagatcgttccatgtcgtatgtgagatga 1191
Db 1378 ttcaataatgtgcatgtgtgtgacgtgctccaaagatcgttccatgtcgtatgtgagatga 1437
QY 1192 gctccccaacggaacattacccctccggaaggtgtcgtgtcggacacacggcgtgtcgaat 1251
Db 1438 gctccccaacggaacattacccctccggaaggtgtcgtgtcggacacacggcgtgtcgaat 1497
QY 1252 gtgtgagacttgtatgacagatcatgttctgtgatcctgtggagatagacataaagta 1311
Db 1498 gtgtgagacttgtatgacagatcatgttctgtgatcctgtggagatagacataaagta 1557
QY 1312 tggacaacaagtaactgtgtgaattgttaagaccttaaatgtgacacaaacggaggtacc 1371

Db 1558 tggacacaaagtaactgttgatgttgaaagacctaataatggaacaaacagagcaattggc 1617
 Qy 1372 tgtttgcaatagaggaagagcgctgagtgatgtgtctatctgcacaactatcagatta 1431
 Db 1618 tgtttgcaatagaggaagagcgctgagtgatgtgtctatctgcacaactatcagatta 1677
 Qy 1432 tggagacataagatgtgtgtcatgttactgagtgatgttgaaagagcgataggaattgtgtcgt 1491
 Db 1678 tggagacataagatgtgtgtcatgttactgagtgatgttgaaagagcgataggaattgtgtcgt 1737
 Qy 1492 tgtatcgtatgtgataaagaagagatgacagtgaggcgctatgataatgaaatlaaagtg 1551
 Db 1738 tgtatcgtatgtgataaagaagagatgacagtgaggcgctatgataatgaaatlaaagtg 1797
 Qy 1552 tggagatcgtgtgagctgttgagaccccgctgtctctgcagaggaacactgtgtctacagac 1611
 Db 1798 tggagatcgtgtgagctgttgagaccccgctgtctctgcagaggaacactgtgtctacagac 1857
 Qy 1612 ctgtgtgagacatccggaagagattttgcactaagatttgatgaattccagattgtcagt 1671
 Db 1858 ctgtgtgagacatccggaagagattttgcactaagatttgatgaattccagattgtcagt 1917
 Qy 1672 agttacatgatgacacaactcctcatctgtggaacttccctaataatgatacagctggccaaagt 1731
 Db 1918 agttacatgatgacacaactcctcatctgtggaacttccctaataatgatacagctggccaaagt 1977
 Qy 1732 gaaccccccgcttcccttcgaaacataacactatcctcagatlaataacatac 1791
 Db 1978 gaaccccccgcttcccttcgaaacataacactatcctcagatlaataacatac 2037
 Qy 1792 tgaactcactactgccaggaaccataaagtgtgcgtatlaac 1836
 Db 2038 tgaactcactactgccaggaaccataaagtgtgcgtatlaac 2082

RESULT 15

AAK51716
 ID AAK51716 standard: cDNA: 2207 BP.

AC AAK51716;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 261.

Human: cytokine; cell proliferation; cell differentiation; gene therapy;

tissue growth factor; immunomodulatory; cancer; Leukemia;

nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001: 2001WO-US04098.

PR 03-FEB-2000: 2000US-0496914.

PR 27-APR-2000: 2000US-0560875.

PR 20-JUN-2000: 2000US-0598075.

PR 19-JUL-2000: 2000US-0620325.

PR 01-SEP-2000: 2000US-0624356.

PR 15-SEP-2000: 2000US-0663561.

PR 20-OCT-2000: 2000US-0693325.

PR 30-NOV-2000: 2000US-0728422.

XX (HSEQ-) HXSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI: 2001-476283/51.
 DR P-PSDB: AAM/8583.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 1: Page 1180-1183; 6221pp; English.
 XX
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM/8323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 SO Sequence 2207 BP: 615 A; 478 C; 555 G; 559 T; 0 other;

Query Match 78.5%; Score 1688.8; DB 22; Length 2207;

Best Local Similarity 94.8%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 67; Indels 30; Gaps 1;

Qy 1 tgcgttgcctgagcgccctggcaccacaagggcgcccgcgagagagcgagccagtgcc 60
 Db 58 tgcgttgcctgagcgccctggcaccacaagggcgcccgcgagagagcgagccagtgcc 117
 Qy 61 tgcgagattatgagcccgccgagcgctgtcgaagagagcgactcaattatgaat 120
 Db 118 tgcgagattatgagcccgccgagcgctgtcgaagagagcgactcaattatgaat 177
 Qy 121 tctcagaagagagaagactgttaataatgagcgaacccctagaagaataacagaaga 180
 Db 178 tctatgccagctctctgt 237
 Qy 181 aatt-----caattagaagacatacaacagct 210
 Db 238 cgtatgctgtataaccagagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 297
 Qy 211 gccagactctgtctaaacccaagaacagatgttttagcaagcactgtatgaagactgag 270
 Db 298 gccagactctgtctaaacccaagaacagatgttttagcaagcactgtatgaagactgag 357
 Qy 271 aattgtgtgccaacaaacaaactgtgccatgtccagtatgtatgtgtgtgtgtgtgtgtgt 330
 Db 358 aattgtgtgccaacaaacaaactgtgccatgtccagtatgtatgtgtgtgtgtgtgtgtgt 417
 Qy 331 cggaaactctcagaagactgtatgaagaagaagagactgtgtgtcaataactttgtggcg 390
 Db 418 cggaaactctcagaagactgtatgaagaagaagagactgtgtgtcaataactttgtggcg 477
 Qy 391 tggatcagagtcagaatcagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 450
 Db 478 tggatcagagtcagaatcagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 537
 Qy 451 caacatgaggaacataactgtatcttaacactatgtgtcagaagaatttcaataactgt 510
 Db 538 caacatgaggaacataactgtatcttaacactatgtgtcagaagaatttcaataactgt 597
 Qy 511 ctgcagctcggagatgtgatacgcctggaagaacattctgcataactgtatgtgcaaa 570
 Db 598 ctgcagctcggagatgtgatacgcctggaagaacattctgcataactgtatgtgcaaa 657
 Qy 571 tcaactatgtgtcgtgaactgt 630
 Db 658 tcaactatgtgtcgtgaactgt 717

QY 631 ctgtgagaagaattatcgagagaatgtgtcagagacagatttctgtgagagagccttgca 690
|||||
Db 718 ctgtgagaagaattatcgagagaatgtgtcagagacagatttctgtgagagagccttgca 777
QY 691 gaacgagaagagatgagagacattatcacaacaaacccctcgagagagatgctct 750
|||||
Db 778 gaacgagaagagatgagagacattatcacaacaaacccctcgagagagatgctct 837
QY 751 cccaactcttltatagagacattatcacaacaaacccctcgagagagatgagacaataa 810
|||||
Db 838 cccaactcttltatagagacattatcacaacaaacccctcgagagagatgagacaataa 897
QY 811 tctaatgtgagatgtgagagacattatcacaacaaacccctcgagagagatgagacaataa 870
|||||
Db 898 tctaatgtgagatgtgagagacattatcacaacaaacccctcgagagagatgagacaataa 957
QY 871 aaagagatttactgttcaagatgtatgatacagaataatgtaagcgccttcgagacaac 930
|||||
Db 958 aaagagatttactgttcaagatgtatgatacagaataatgtaagcgccttcgagacaac 1017
QY 931 acaataaagatctggagataaaacacattggaatgcgaacgaattctcacaagccataca 990
|||||
Db 1018 acaataaagatctggagataaaacacattggaatgcgaacgaattctcacaagccataca 1077
QY 991 ggtcagctcctgtctcagatgtatgagagagatgatacacaagatcagatcc 1050
|||||
Db 1078 ggtcagctcctgtctcagatgtatgagagagatgatacacaagatcagatcc 1137
QY 1051 acggtcagagtggtgagatgtaaatatagatgtaaatgtaaatgtaaatgtaaatgta 1110
|||||
Db 1138 acggtcagagtggtgagatgtaaatatagatgtaaatgtaaatgtaaatgtaaatgta 1197
QY 1111 gaagcagattctgagactgtgcttcaataatgagatgtatgagcgtcccaagatcgt 1170
|||||
Db 1198 gaagcagattctgagactgtgcttcaataatgagatgtatgagcgtcccaagatcgt 1257
QY 1171 tccatgtctgtatgagatgtgagcctcccaactgacattacccctcgagagtgctgtc 1230
|||||
Db 1258 tccatgtctgtatgagatgtgagcctcccaactgacattacccctcgagagtgctgtc 1317
QY 1231 ggaacccgagctgtctcaatgtgtgagacttgatgacaagtacattgttctgcatct 1290
|||||
Db 1318 ggaacccgagctgtctcaatgtgtgagacttgatgacaagtacattgttctgcatct 1377
QY 1291 ggggagatgaactataaaggtatggaacacacagactgtgaaattgttaaggaacttaaat 1350
|||||
Db 1378 ggggagatgaactataaaggtatggaacacacagactgtgaaattgttaaggaacttaaat 1437
QY 1351 ggaacacaaagagcattgctgtgttcagttacagagagagcgtgtgtagtgcgtca 1410
|||||
Db 1438 ggaacacaaagagcattgctgtgttcagttacagagagagcgtgtgtagtgcgtca 1497
QY 1411 tctgacaacactatcagatataatgagacataagatgtgtgtcagatgttagaa 1470
|||||
Db 1498 tctgacaacactatcagatataatgagacataagatgtgtgtcagatgttagaa 1557
QY 1471 gggcabaaggaattgtgtcgtgtatcagattgatacaacaagagataatgtagtgcgtcc 1530
|||||
Db 1558 gggcabaaggaattgtgtcgtgtatcagattgatacaacaagagataatgtagtgcgtcc 1617
QY 1531 tatgtatgaaaataatlaaagtgtgagatctgtgtcgtcttgagaccctgtccttgca 1590
|||||
Db 1618 tatgtatgaaaataatlaaagtgtgagatctgtgtcgtcttgagaccctgtccttgca 1677
QY 1591 gggacactctgtctacgagacctgtgtgagagcattccggaagagtttltgactacagttt 1650
|||||
Db 1678 gggacactctgtctacgagacctgtgtgagagcattccggaagagtttltgactacagttt 1737
QY 1651 gatgaattccagattgtcagatgtatcacaatgatacacaatccctcaatctgagacttccca 1710
|||||
Db 1738 gatgaattccagattgtcagatgtatcacaatgatacacaatccctcaatctgagacttccca 1797

QY 1711 aatgatccagctgcccacagctggaaccccccggttcccttctcgaacatacacaactacac 1770
|||||
Db 1798 aatgatccagctgcccacagctggaaccccccggttcccttctcgaacatacacaactacac 1857
QY 1771 tccagataaataacacatacactgaactgaactgtcccgagacccttaagtgtgagta 1830
|||||
Db 1858 tccagataaataacacatacactgaactgaactgtcccgagacccttaagtgtgagta 1917
QY 1831 tttaac 1836
|||||
Db 1918 cataac 1923

Search completed: May 11, 2002, 09:23:17
Job time: 11031 sec

TITLE
EST (Bloeker, et al.)
JOURNAL
Unpublished (1999)

CTTTTGTGATGATTGGACCTTTAAACCTCCCTCCTCCTCCTTTCACCTCTGCA 598

CTTTTGTAATGATTGGACTTTAAACCTCCCTCTCTCTCTTTCACCTTGCA 598

| | | | | |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match | 29.5%; | Score 635.2; | DB 9; | Length 652; |
| Best Local Similarity | 99.2%; | Pred. No. 1.4e-164; | | |
| Matches 649; | Conservative 0; | Mismatches 3; | Indels 2; | Gaps 1; |

Best Local Similarity 96.3%; Pred. No. 1,8e-158;
Matches 649; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

```

Oy 602 aatgtaacagagtgacccctgtagatgcatgctgtagaagaactatctgagaagatgta 661
    |||
Db 1 AATGTAACAGAGTGAACCTGTGATGCGATGCTGTGAAACAACTTATCCAGAGAAATGTC 60
Oy 662 gacagatctctgtgagagagccctgacagacgaagagatgggagacagatattatca 721
    |||
Db 61 GGACAGATTCTCTGTGGAGAGGCCCTGGACAGAACAGAGATGGGACAGATTTATTATCA 120
Oy 722 aaacaaacctctgagcgggaatgctcctccaaactcttttataagacattatccta 781
    |||
Db 121 AAAACAAACCTCTGACCGGAGATGCTCTCCCACTCTTTTATAGACACTTATCTCTA 180
Oy 782 aaattatcaagacattgagacaaatagaatctaatgtgagatgtggaagacatggttac 841
    |||
Db 181 AATATTATCAAGACATTTGAGACAAATAGATCTAATTGAGATGTGGAAGACATAGTTTAC 240
Oy 842 agagaattcactgcccgaagtgaacaaagcaagagattactgtttacagatgatgac 901
    |||
Db 241 AGAGAAATTCACCTGCCGAGTGAACCAAGAGATTACTGTTTACATATCATATC 300
Oy 902 agaaatagttaagggccttcgagacaacacatcaagaatctgggataaaacacattg 961
    |||
Db 301 AGAAATATGTAAGCGCCCTTCGAGACAAACAAATCAAGATCTGGGATAAAACACATTCG 360
Oy 962 aatgcaacagatctctcaacagacatagatcagatcctctgtctcagatgatgaga 1021
    |||
Db 361 AATGCAACGAAATTCATCAGCCATACAGGATTCAGTCTCTGCTCCAGATATCATAGA 420
Oy 1022 gatgatacatacagagatcatcgatccagtcagagtgtagagatg-aaatacagat 1080
    |||
Db 421 GAGTATCATTAACAGGATCATCGGATTCACAGGTGAGGTGAGATGTAATATACAGT 480
Oy 1081 gaatgcaaacacgattgattaccattgtgaacagatctgcaccttgcgcttcaataat 1140
    |||
Db 481 GAAATGCTAAACAGTTGATTCACCATTTGTAACAGGTGCACTTGCGCTTCATTAATAT 540
Oy 1141 ggcattgagtgagacgtctcccaagaatcgttccatctgtatgagatagagcctccca 1200
    |||
Db 541 GGCATGATGTGACCTGCTCCAAAGATCGTTCATGCTGATGAGATGATGCTCCCCCA 600
Oy 1201 actgacattaccctccggaaggtgctgctc-ggacaccgaagctgctgtaaatgttga 1259
    |||
Db 601 CTGACATTAACCTCCGAGAGGTGCTGTCGGACCAAGGCTGTGATCATGTTTGA 660
Oy 1260 cttgatgacaagt 1273
    |||
Db 661 CTTGCTTGGCGGT 674

```

RESULT 5
BE269297
LOCUS 692 bp mRNA linear EST 13-JUL-2000
DEFINITION 601186356F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3544549 5',
mRNA sequence.
ACCESSION BE269297
VERSION BE269297.1 GI:9142915
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gqabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LNCM239 Row: 0 Column: 14
High quality sequence stop: 610.
Location/Qualifiers

FEATURES

source

1..692

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3544549"

/clone_11b="NIH_MGC_8"

/tissue_type="Burkitt lymphoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph. Vector: pCMV7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

Db 701 AATGTCAGACAGATTCTCTGTGAGAGGCGCTGGCAGAGAACAAG 748

RESULT 9
LOCUS BG518761
DEFINITION 602578346f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491843 5', mRNA sequence.

ACCESSION BG518761
VERSION BG518761.1 GI:13513945
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 789)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM8536 row: k column: 12
High quality sequence stop: 587.

FEATURES
source
1..789
/organism="Mus musculus"
/strain="C57BL/6J (f1)"
/db_xref="taxon:10090"
/clone="IMAGE:3491843"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary."
/stem_cell_origin=""
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site: 1; SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 257 a 176 c 202 g 154 t

ORIGIN

Query Match 24.88; Score 534.4; DB 10; Length 789;
Best Local Similarity 86.18; Pred. No. 1.2e-136;
Matches 661; Conservative 0; Mismatches 101; Indels 6; Gaps 6;

65 cgaatagagccgagccgagcggtgctgaagaagagcaccgaattatgaatcct 124
|||||
1 CGATTATGAGACCGCGAGCGCGTGTGACAGAGAAAGCGCTTAATATATGAATTCCT 60
|||||

125 cagagagagaagactgtaataatgagcaaccocctaggaagataatcacagagaagact 184
|||||
61 CAGAGAGAGAAGACTGTAATATATGCGAACCCTAGAGAAATATACCGAGAGAAGATT 120
|||||

185 cacttaagagacatacaacagctgctgcagactctgcttaaaccaagaacagatggt 244
|||||
121 CACTTAAACAGACTTAAACAAGCTGTGCGCAGGCTTTGCTATTAACCAAGACAGTATGTC 180
|||||

245 taagaagcactgctataaagaactgagaaatgctgagcacaacaaactgccaatggca 304
|||||
181 TAAACAAGCAGCTATGATGAGACTGAATAATGTGTGGCCAAAGCAAACTTGCCAAATGGCA 240
|||||

305 ctccagatagatgctgcccagaacagaaactctcagcagcagctatgaaagaaag 363
|||||
241 CTTCACAGATGATGTCGCCCAAGCGGAAACCTCAGCAGAGGTATGAGAGAGAACAG 300
|||||

364 gaactggtgcaataacttgagcagtggtcagagtcagatcaagtgg-aatttggga 422

Db 301 GAGCTGTGTCAAGTATTTTGGACAGGTGTCACAGTCTGATCAAGTGAATTTGTAGA 360
|||||

423 acatctatcccaaatgltgltacatcccaatggtgcaataaactcgtatctaaacc 482
|||||

361 AACCTTATATCCCAAAAGTGTACTACACAGATGGCGACATCAACTCACTTAACCAACC 420
|||||

483 tatgttgagaagagatttcataactgctctgcccagccctcggagatgatgatcgtgta 542
|||||

421 TATGCTGAGAGGAGATTTTCATTAAGTGCACCTGCGACAGGGGCTGTGACCAATCGCTGA 480
|||||

543 gaacatctgcatatccctgagatgccaatcactatgctgctgaactgtgtgcaagga 602
|||||

461 GACATTTCTGTCATCTTGAGACGCCAACTCACTGTCTCTCTGATGCTGTGCAAGGA 540
|||||

603 atgtaccg-agtgaacctgtgagcagctgtggaagaagcttatcagagaagtgtca 661
|||||

541 ATGTACCGCAGAGTACGTCGCGACGATGCTGTGAAAACCTCATCGACAGATGCTGA 600
|||||

662 ggaac-agaattctgtgga-gagcctgtgcaagcagaagagatggagacgtattatc 719
|||||

601 GACCGGCTACTCTGTGACAGAGCGCCGACAGCGCTGGGACGACTTACTC 660
|||||

720 caaacaacacctctgagcggaatgctccccaactctttatagagacattatcc 779
|||||

661 AAAACCAAACTCATGATGAGCAAGCGCTCTCCAAAGAGCAATAAGACGGCTTAACC 720
|||||

780 taatatatacaagacattgagacaatagaatcctaattgagagatggtg 827
|||||

721 TAAATCATTA-TAAGACATAGAGACAAATTAAGTCAATGAGAAATGGGG 767
|||||

RESULT 10
LOCUS B1103092 637 bp mRNA linear EST 26-JUN-2001
DEFINITION 602889321P1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044684
5', mRNA sequence.

ACCESSION B1103092
VERSION B1103092.1 GI:14553985
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 637)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM11122 row: i column: 05
High quality sequence stop: 629.

FEATURES
source
1..637
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5044684"
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library. I"

BASE COUNT 185 a 148 c 172 g 132 t

VERSION B145545.1 GI:15246201
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 888)
 NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-riemail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINT at:
<http://image.lnl.gov>
 Plate: L1639 row: d column: 19
 High quality sequence start: 5
 High quality sequence stop: 684.
 Location/Qualifiers

FEATURES

source

1..888

/organism="Mus musculus"

/strain="C57/B6"

/db_xref="taxon:10090"

/clone_image="5253090"

/clone_lib="NCI-CGAP_Mam5"

/tissue_type="tumor, gross tissue"

/dev_stage="7 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1: SalI;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"

BASE COUNT 219 a 218 c 248 g 203 t
 ORIGIN

Query Match 24.6%; Score 528.4; DB 10; Length 888;
 Best Local Similarity 81.6%; Pred. No. 6e-135;
 Matches 725; Conservative 0; Mismatches 141; Indels 22; Gaps 9;

QY 854 gccgaagatgaacaaagagattactgtttacagatgatgcagaataatgtaa 913
 DB 1 gccgaagatgaacaaagagattactgtttacagatgatgcagaataatgtaa 913
 QY 914 gcggccttcgagacacacacatcagatctggatataaacaacatgtaacagaa 973
 DB 61 gcggccttcgagacacacacatcagatctggatataaacaacatgtaacagaa 973
 QY 974 ttctacagagcatagagttcagtcctctgtctcagatgatgcagaataatgtaa 1033
 DB 121 ttctacagagcatagagttcagtcctctgtctcagatgatgcagaataatgtaa 1033
 QY 1034 cagagatcagatccacagatcagatggtgagatgaataacatgtaacagaa 1093
 DB 181 cagagatcagatccacagatcagatggtgagatgaataacatgtaacagaa 1093
 QY 1094 cgttattcaccatgtgtaa-gcagttctgacatcgcttcaataatgtaacagaa 1152
 DB 241 cgttattcaccatgtgtaa-gcagttctgacatcgcttcaataatgtaacagaa 1152
 QY 1153 accgttcacaagaatcgttccatgtglatggaatagtcctcccaactgatac 1212
 DB 301 accgttcacaagaatcgttccatgtglatggaatagtcctcccaactgatac 1212
 QY 1213 ctccgagaggtgctgtgagaaaccgagctgctcaatggttgatgataacag 1272
 DB 361 ctccgagaggtgctgtgagaaaccgagctgctcaatggttgatgataacag 1272

QY 1273 tacatgtttctgcacatctgggagtagaactataaagatgatgacaacaactgtgtgaa 1332
 DB 421 tacatgtttctgcacatctgggagtagaactataaagatgatgacaacaactgtgtgaa 1332
 QY 1333 ttgttagaagacctaataatgagacaacagagatgctggtttgagatgaagagcagg 1392
 DB 481 ttgttagaagacctaataatgagacaacagagatgctggtttgagatgaagagcagg 1392
 QY 1393 ctgtgtagtgcacatctgacacacatcagatgataatgagacatag-aatgtgtgc 1451
 DB 541 ctgtgtagtgcacatctgacacacatcagatgataatgagacatag-aatgtgtgc 1451
 QY 1452 atgttaacagatgttgaagagcacaagaagaa-atgtgtgtgtatcattgataac- 1509
 DB 601 atgttaacagatgttgaagagcacaagaagaa-atgtgtgtgtatcattgataac- 1509
 QY 1510 aagagatagtcagtggtggcctatgataaataaagatgtggatc-----ttgtg 1564
 DB 661 aagagatagtcagtggtggcctatgataaataaagatgtggatc-----ttgtg 1564
 QY 1565 ctgcttgaaccccgctgctcctcagagagacactgtctac-----gaacctgtgag 1620
 DB 721 ctgcttgaaccccgctgctcctcagagagacactgtctac-----gaacctgtgag 1620
 QY 1621 c---atccggaagatgttgcactac-----agttgatgaattccag-atgtcagt 1671
 DB 781 c---atccggaagatgttgcactac-----agttgatgaattccag-atgtcagt 1671
 QY 1672 agtccatgatgacacacatcctcactcagtgagacttccatgatcaca 1719
 DB 841 agtccatgatgacacacatcctcactcagtgagacttccatgatcaca 1719

RESULT 13
 LOCUS BG082149 646 bp mRNA linear EST 26-JAN-2001
 DEFINITION H3073A06-5 N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 accession BG082149
 VERSION BG082149.1 GI:12564717
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 646)
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
 /T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of N1A mouse 15K cDNA clone set
 Unpublished (2001)
 Other ESTs: H3073A06-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@jgk.sun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://lgsun.grc.nia.nih.gov/cDNA/15K.html> for details.
 Plate: H3073 row: A column: 06
 Seq primer: -21M13 Reverse
 High quality sequence stop: 646
 POLYA-No.

FEATURES

source

Location/Qualifiers

1..646

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone_image="5253090"

/clone_lib="NCI-CGAP_Mam5"

/tissue_type="tumor, gross tissue"

/dev_stage="7 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1: SalI;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"

```

/lab host="DH10B"
/Note="Vector: pSPORT1; Site.1: Salt; Site.2: NotI; This
clone is among a rearranged set of 15,247 clones from 11
embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5 embryos
, and E12.5 female mesonephros/gonad) and one newborn
ovary cDNA library. Average insert size 1.5 kb. All
source libraries are cloned unidirectionally with Oligo(dT
) Not primers. References include: (1) Genome-wide
expression profiling of mid-gestation placenta and embryo
using a 15,000 mouse developmental cDNA microarray, 2000,
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
Large-scale cDNA analysis reveals phased gene expression
patterns during preimplantation mouse development, 2000,
Development, 127: 1737-1749; (3) Genome-wide mapping of
unselected transcripts from extraembryonic tissue of
7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the X chromosome, 1998, Hum
Mol Genet 7: 1967-1978."
Mol Genet 7: 1967-1978."

```

```

BASE COUNT      150 a      149 c      198 g      149 t
ORIGIN
Query Match      23.3%; Score 500.4; DB 10; Length 646;
Best Local Similarity 87.2%; Pred. No. 2.9e-127;
Matches 560; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

```

```

QY 982 ggcacatacaggttcagtcctctgtccagatgatgagagatgatacacaagatca 1041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GGGCAGACAGCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
QY 1042 tggatattccaggttcagatgagatgagatgagatgagatgagatgagatgagat 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TCAGACTCCACCGCTCAGAGTGTGGGATGTAAATGCAAGTGTGATGCTAAACACTTATT 121
QY 1102 caccattgtgaagcagcttgcacttgcacttgcacttgcacttgcacttgcacttgc 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 CACCACTGTGAAGCGCTTTCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
QY 1162 aagatcgttccatgtgtgatgataatgagatgagatgagatgagatgagatgagat 1221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 AAAGACCGCTTCCATGCTGTGTGGGATGTGGCTTCCCACTGACATCACCCTTAGAGG 241
QY 1222 gtgtgtgtgacacagcagctgtcacttgcacttgcacttgcacttgcacttgcact 1281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 GTCTGTGTGGGACACCGAGCTGCGGTCAATGTTGTAGACTTTGATGACAGTACATGCTT 301
QY 1282 tctgcatctgggagatagaactataaaggtatgaaacacagacttgaatt-tgtaag 1340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TCTGCTCTGAGATGAACCATTAAGTGGGCAACACAGAGACCTGGAAATTCGGAAG 361
QY 1341 gaccttaaatgagacacaaagagcattgctgtgttgacagacagcagcagcagcagc 1400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 GCCCTTAAATGGGACAGAGGTGGCATCGCTGTTTGCAGTACGAACAGACAGCGTGGTGT 421
QY 1401 gagtgtctcttgcacacacactacatgatgatgagacatgaagtgtgtgtgtgtgtgt 1460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 GAGGCGCTCTCTGACAAACACAGCTGTGGGACATGAGTGTGGAGATGCCCTGCG 481
QY 1461 agtgttagaagcactgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 AGTGTGTGAGAGGCGCATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 541
QY 1521 cagtgggagcctatgatgagaaataaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 GAGGAGGAGCCTATGATGGGAAATTAAGTGTGGGATCTTATGAGCTCTTGGAGCCGCG 601
QY 1581 tgtctcctgagagacacttcttctacgacaccttgttgagca 1622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 TGCTCCAGCAGGAGCTCTCTGTCTGCGGACACTTGTGAAGCA 643

```

```

RESULT 14
BG058605/c 541 bp mRNA linear EST 25-JAN-2001
DEFINITION
naflieh04.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:411302 3'
SIMILAR to TR:09Y213 09Y213 BETA-TRANSDUCIN REPEATS CONTAINING
PROTEIN. ; mRNA sequence.
ACCESSION
BG058605
VERSION
BG058605.1 GI:12525270
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 541)
NCl-CGAP http://www.ncbi.nlm.nih.gov/cgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: M. Bento Soares, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima
Bonafide, Ph.D.

```

```

COMMENT
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCl-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
infoimage.lnl.gov
Seq primer: -400P from Glibco
High quality sequence stop: 441.
Location/Qualifiers
1. 541
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:411302"
/clone_lib="Soares_NPBMC"
/tissue_type="Lymphocyte"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: blood; Vector: pTZ19-D-Pac; Site_1: NotI;
Site_2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTCACCAATCTGAAAGTGGGAGCGCGCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pTZ19 vector. Library
is normalized; constructed in the laboratory of M. Bento
Soares (University of Iowa)."

```

FEATURES

source

```

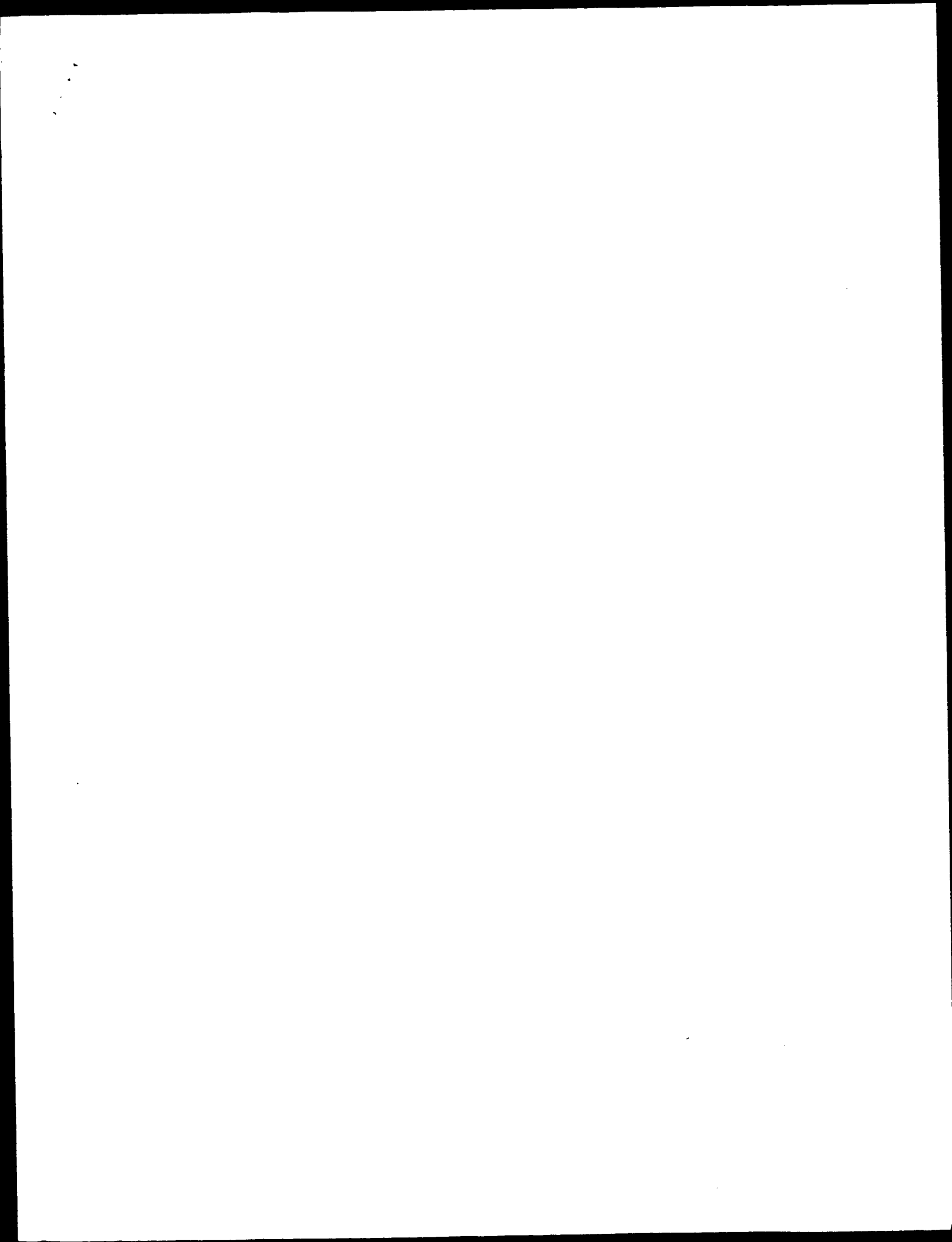
BASE COUNT      140 a      110 c      145 g      146 t
ORIGIN
Query Match      22.9%; Score 492.2; DB 10; Length 541;
Best Local Similarity 95.6%; Pred. No. 5e-125;
Matches 517; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

```

```

QY 1603 ctacgagccttgtgagcattccggaagatttgcactacagtttgatgaattcag 1662
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CTACGAAATCAATTTGGAGCAATTTCTGAAGATTATGACACACAGTTTGAAGAAGCCAG 482
QY 1663 attgtcgttgcacttgcacttgcacttgcacttgcacttgcacttgcacttgcact 1722
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 AGTGGCAGAAATTGACATGATGACACAAATCCGACGCGGGACCTCTTAATGATCCAGCA 422
QY 1723 gcccaagctgaaccccccttcccttccgagacatacacactacatcctcaagataata 1782
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GCACAACTGAACCCGCCCTTCCCTTCTGGAACAACACACTCAATCTCCAGATAATA 362
QY 1783 accatacctgacactacttgcacagacacacacacacacacacacacacacacacacac 1842
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 ACCATACACTGACCTCAAACTTGGCCAGACCCCTTAATGTTGGGATTTAAGATATCT 302
QY 1843 gccatcacgagatgagacacacacacacacacacacacacacacacacacacacacac 1902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:49:58 ; Search time 28.99 Seconds

(without alignments)
2870.954 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034

Sequence: 1 MDPAAVLAQKALKFNNSS.....PAAQAPPPSPRTYIISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :
1: SPREMBL_17.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp Vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|--------------------|
| 1 | 2997 | 98.8 | 569 | 11 090U15 | Q9qu15 mus musculu |
| 2 | 2990 | 98.5 | 569 | 11 092159 | Q92159 mus musculu |
| 3 | 2978 | 98.2 | 569 | 11 09R1G7 | Q9r1g7 mus musculu |
| 4 | 2089.5 | 68.9 | 510 | 5 044382 | 044382 drosophila |
| 5 | 2089.5 | 68.9 | 510 | 5 09VDE3 | 09vde3 drosophila |
| 6 | 1638.5 | 54.0 | 665 | 5 09GNN6 | 09gnn6 caenorhabd |
| 7 | 1498 | 49.4 | 430 | 5 09B554 | 09b554 heterodera |
| 8 | 690 | 22.7 | 506 | 3 09P7V1 | 09p7v1 schizosacch |
| 9 | 661 | 21.8 | 533 | 5 09V2F4 | 09v2f4 drosophila |
| 10 | 640 | 21.1 | 553 | 5 09NUX6 | 09nux6 homo sapien |
| 11 | 520 | 17.1 | 587 | 5 044083 | 044083 caenorhabd |
| 12 | 387.5 | 12.8 | 304 | 11 09P4T2 | 09p4t2 mus musculu |
| 13 | 365.5 | 12.0 | 942 | 5 096611 | 096611 dictyostell |
| 14 | 356 | 11.7 | 410 | 13 09P7R5 | 09p7r5 gallus gall |
| 15 | 351 | 11.6 | 410 | 6 09GL51 | 09gl51 sus scrofa |
| 16 | 344 | 11.3 | 411 | 5 096698 | 096698 drosophila |
| 17 | 339.5 | 11.0 | 777 | 3 09USN3 | 09usn3 schizosacch |
| 18 | 334.5 | 11.0 | 283 | 11 09JUV3 | 09juv3 mus musculu |
| 19 | 334.5 | 11.0 | 454 | 4 09HA09 | Q9ha09 homo sapien |

| | | | | | |
|----|-------|------|------|-----------|--------------------|
| 20 | 332.5 | 11.0 | 317 | 10 09M2Z2 | Q9m2z2 arabidopsis |
| 21 | 332 | 10.9 | 454 | 11 09QUH1 | Q9quh1 mus musculu |
| 22 | 330.5 | 10.9 | 594 | 4 09H073 | Q9h073 homo sapien |
| 23 | 330 | 10.9 | 333 | 10 09SY00 | Q9sy00 arabidopsis |
| 24 | 325 | 10.7 | 277 | 11 09R2A6 | Q9r2a6 mus musculu |
| 25 | 325 | 10.7 | 391 | 11 035592 | 035592 mus musculu |
| 26 | 324.5 | 10.7 | 328 | 11 09D7H2 | Q9d7h2 mus musculu |
| 27 | 324.5 | 10.7 | 334 | 4 09NMK7 | Q9nmk7 homo sapien |
| 28 | 313.5 | 10.3 | 481 | 5 09VPR4 | Q9vpr4 drosophila |
| 29 | 313.5 | 10.3 | 485 | 4 09NVX2 | Q9nvx2 homo sapien |
| 30 | 313.5 | 10.3 | 487 | 4 09BU54 | Q9bu54 homo sapien |
| 31 | 312.5 | 10.3 | 330 | 4 09NU14 | Q9nu14 homo sapien |
| 32 | 312 | 10.3 | 876 | 10 09LFE2 | Q9lfe2 arabidopsis |
| 33 | 308 | 10.2 | 480 | 5 096995 | 096995 drosophila |
| 34 | 304.5 | 10.0 | 1205 | 4 09U666 | Q9u666 homo sapien |
| 35 | 304.5 | 10.0 | 1205 | 4 09U599 | Q9u599 homo sapien |
| 36 | 301.5 | 9.9 | 476 | 13 093531 | Q93531 xenopus lae |
| 37 | 299.5 | 9.9 | 1205 | 4 09U655 | Q9u655 homo sapien |
| 38 | 298.5 | 9.8 | 1171 | 4 09U558 | Q9u558 homo sapien |
| 39 | 298.5 | 9.8 | 1205 | 4 09U677 | Q9u677 homo sapien |
| 40 | 298.5 | 9.8 | 1205 | 4 09U654 | Q9u654 homo sapien |
| 41 | 298.5 | 9.8 | 1205 | 4 09U633 | Q9u633 homo sapien |
| 42 | 297 | 9.8 | 502 | 3 074855 | 074855 schizosacch |
| 43 | 295.5 | 9.7 | 1205 | 4 09U650 | Q9u650 homo sapien |
| 44 | 294.5 | 9.7 | 1249 | 11 09EPV5 | Q9epv5 rattus norv |
| 45 | 293.5 | 9.7 | 1205 | 4 09U652 | Q9u652 homo sapien |

ALIGNMENTS

RESULT 1
ID Q90U15 PRELIMINARY; PRT; 569 AA.
AC Q90U15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE UBQUITIN LIGASE FWD1 (BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN).
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99199275; PubMed=10097128;
RA Hakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-I.,
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase.";
RT Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Strausberg R.,
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF081887; AAD17755.1; -
DR EMBL: AF099932; AAD08701.1; -
DR EMBL: BC003989; AA03989.1; -
DR MGD: MGI:1338671; Btrc.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001880; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.

PRINTS: PR00320; GPROTEINBPT.
 DR SMART: SM00256; FBOX: 1.
 DR SMART: SM00320; WD40: 7.
 DR PROSITE: PS50181; FBOX: 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ligase; Repeat; WD repeat.
 SO SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;
 Best Local Similarity 98.6%; Pred. No. 5e-239;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCINQETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCINQETVCLT 60
 QY 61 STAMKTENCVAKTLANGTSSMIVPQKRLASYEKEKELCYFFQWESDQVEVEHL 120
 DB 61 STAMKTENCVAKAKLANGTSSMIVPQKRLASYEKEKELCYFFQWESDQVEVEHL 120
 QY 121 ISQCHYOHGHINSYIKPMQDPITLAPRGLDHIAENILSYLDAKSICAAELVCKEM 180
 DB 121 ISQCHYOHGHINSYIKPMQDPITLAPRGLDHIAENILSYLDAKSICAAELVCKEM 180
 QY 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDENAPNSFYRALYPKII 240
 DB 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDENAPNSFYRALYPKII 240
 QY 241 QDIETIESNWRGCRHSIORHCRSETSGVYCYQYDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNWRGCRHSIORHCRSETSGVYCYQYDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKIVNTSTC 420
 DB 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKIVNTSTC 420
 QY 421 EFVRLTNGHKGRIACLOQYDRDLVYSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDN 480
 DB 421 EFVRLTNGHKGRIACLOQYDRDLVYSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDN 480
 QY 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGVFRLOPDEFOIVSSSHDT 540
 DB 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGVFRLOPDEFOIVSSSHDT 540
 QY 541 IILWFLNDPAAQAEPSPSRITYYISR 569
 DB 541 IILWFLNDPAAQAEPSPSRITYYISR 569

RESULT 2
 Q92159 PRELIMINARY; PRT; 569 AA.
 AC Q92159;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
 GN BTRC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99145465; PubMed=9990853;
 RA Spencer E., Jiang J., Chen Z.J.;

RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
 RT SLIMB/Beta-TrCP.";
 RL Genes Dev. 13:284-294(1999).
 DR EMBL: AF112979; AAD04181.1; -.
 DR MGI: 1338871; Btrc.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBPT.
 DR SMART: SM00256; FBOX: 1.
 DR SMART: SM00320; WD40: 7.
 DR PROSITE: PS50181; FBOX: 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SO SEQUENCE 569 AA; 65047 MW; BC7CA44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;
 Best Local Similarity 98.4%; Pred. No. 1.9e-238;
 Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCINQETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCINQETVCLT 60
 QY 61 STAMKTENCVAKTLANGTSSMIVPQKRLASYEKEKELCYFFQWESDQVEVEHL 120
 DB 61 STAMKTENCVAKAKLANGTSSMIVPQKRLASYEKEKELCYFFQWESDQVEVEHL 120
 QY 121 ISQCHYOHGHINSYIKPMQDPITLAPRGLDHIAENILSYLDAKSICAAELVCKEM 180
 DB 121 ISQCHYOHGHINSYIKPMQDPITLAPRGLDHIAENILSYLDAKSICAAELVCKEM 180
 QY 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDENAPNSFYRALYPKII 240
 DB 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDENAPNSFYRALYPKII 240
 QY 241 QDIETIESNWRGCRHSIORHCRSETSGVYCYQYDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNWRGCRHSIORHCRSETSGVYCYQYDQKIVSGLRDNTIKIMDKNTLECK 300
 QY 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RILTGHTGSVLCLOYDERVITIGSSDSYVRWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKIVNTSTC 420
 DB 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKIVNTSTC 420
 QY 421 EFVRLTNGHKGRIACLOQYDRDLVYSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDN 480
 DB 421 EFVRLTNGHKGRIACLOQYDRDLVYSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDN 480
 QY 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGVFRLOPDEFOIVSSSHDT 540
 DB 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGVFRLOPDEFOIVSSSHDT 540
 QY 541 IILWFLNDPAAQAEPSPSRITYYISR 569
 DB 541 IILWFLNDPAAQAEPSPSRITYYISR 569

RESULT 3
 Q9RIG7 PRELIMINARY; PRT; 569 AA.
 AC Q9RIG7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF110396; AAd41023.1; -
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D51D9D CRC64;

Query Match 98.2%; Score 2978; DB 11; Length 569;

Best Local Similarity 97.9%; Pred. No. 1.9e-237;

Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MDPAAEVLQEKALKFNSSEREDCNNGEPKRIPEKNSLRQTSYSCARLCINQETVCLIA 60
DB 1 MDPAAEVLQEKALKFNSSEREDCNNGEPKRIPEKNSLRQTSYSCARLCINQETVCLIT 60
QY 61 STRAKTENCVAATKLANGTSSMIVPKOKRLSAYEKEKELCVKYEQWSESQVFEVHL 120
DB 61 STRAKTENCVAATKLANGTSSMIVPKOKRLSAYEKEKELCVKYEQWSESQVFEVHL 120
QY 121 ISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKEY 180
DB 121 ISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKEY 180
QY 121 ISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKEY 180
DB 121 ISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKEY 180
QY 181 RYTSDDMLKKLIERVNRDTSIMRGLAERGGQYLFKNKPPDGNAPPSFYRALYPII 240
DB 181 RYTSDDMLKKLIERVNRDTSIMRGLAERGGQYLFKNKPPDGNAPPSFYRALYPII 240
QY 241 QDIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNDNTIKIWDKNTLECK 300
DB 241 QDIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNDNTIKIWDKNTLECK 300
QY 301 RILTGHTGSLVLCQYDERIITIGSSDSITYRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 360
DB 301 RILTGHTGSLVLCQYDERIITIGSSDSITYRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 360
QY 361 VYCSKDRSAIVMDMASPTDITLRVLYGHRAAVNVVDFDCKIYASASGDRITKWNSTSC 420
DB 361 VYCSKDRSAIVMDMASPTDITLRVLYGHRAAVNVVDFDCKIYASASGDRITKWNSTSC 420
QY 421 EYVRLNGHKGRIACIQYDRDLVYSGSSDNTIRLMDIECGACRLVLEGHLELVRCIRFN 480
DB 421 EYVRLNGHKGRIACIQYDRDLVYSGSSDNTIRLMDIECGACRLVLEGHLELVRCIRFN 480
QY 481 KRIVSAGYGGIKVMDLVVALDPRAPAGTLCRTLVEHSGVFRLOPDEFOIVSSSHDT 540
DB 481 KRIVSAGYGGIKVMDLVVALDPRAPAGTLCRTLVEHSGVFRLOPDEFOIVSSSHDT 540
QY 541 ILIWDPLNDPAQAEPSPSRITYISR 569
DB 541 ILIWDPLNDPAQAEPSPSRITYISR 569

RESULT 4
ID 044382 PRELIMINARY: PRT: 510 AA.
AC 044382;

DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SLIMB.
GN SLIMB OR SLIMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9812115; PubMed-9461217;
RA Jiang J., Struhl G.;
RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb.";
RL Nature 391:493-496(1998).
DR EMBL; AF032878; AAC3852.1; -
DR FlyBase; FBgn0023423; slimb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;

Best Local Similarity 78.3%; Pred. No. 3.4e-164;

Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;

QY 64 MTEKNCVAKTKLANG---TSSMIVPKOKRLSAS--YEKEKELCVKYEQWSESQVFEV 118
DB 4 MTEKNCVAKTKLANG---TSSMIVPKOKRLSAS--YEKEKELCVKYEQWSESQVFEV 118
QY 119 HLISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKE 178
DB 63 HLISOMCHYOHGHNISYKPMLODFITLAPARGLDHAENILSYLAKSICAEELVCKE 178
QY 179 WYRYSDDMLKKLIERVNRDTSIMRGLAERGGQYLFKNKPPDGNAPPSFYRALYPII 238
DB 123 WYRYSDDMLKKLIERVNRDTSIMRGLAERGGQYLFKNKPPDGNAPPSFYRALYPII 238
QY 239 IIDIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNDNTIKIWDKNTLECK 298
DB 182 IIDIETIESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIVSGLDNDNTIKIWDKNTLECK 298
QY 299 KRLITGHTGSLVLCQYDERIITIGSSDSITYRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 358
DB 242 KRLITGHTGSLVLCQYDERIITIGSSDSITYRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 358
QY 359 MMYTCSKDRSAIVMDMASPTDITLRVLYGHRAAVNVVDFDCKIYASASGDRITKWNSTSC 418
DB 302 MMYTCSKDRSAIVMDMASPTDITLRVLYGHRAAVNVVDFDCKIYASASGDRITKWNSTSC 418
QY 419 TCEYVRLNGHKGRIACIQYDRDLVYSGSSDNTIRLMDIECGACRLVLEGHLELVRCIRF 478
DB 362 TCEYVRLNGHKGRIACIQYDRDLVYSGSSDNTIRLMDIECGACRLVLEGHLELVRCIRF 478
QY 479 DNKRIVSAGYGGIKVMDLVVALDPRAPAGTLCRTLVEHSGVFRLOPDEFOIVSSSHD 538
DB 422 DNKRIVSAGYGGIKVMDLVVALDPRAPAGTLCRTLVEHSGVFRLOPDEFOIVSSSHD 538
QY 539 DTILWDPLNDPAQAEPSPSRITYISR 561
DB 482 DTILWDPLNDPAQAEPSPSRITYISR 561

[illegible]

DB 1181 DYTGNCKHTLNGHOSLTSGLMELRONILVSGNADSTVAKWIDITG-----OCLQTL 1231
 QY 516 V---EHSGRVRLQDEFOIVSSSHDDITLWD 545
 DB 1232 SGPKNHSAVTLCLQFNRSFVTSDDGTIVKLWD 1264

RESULT 10

QY 09N0X6 PRELIMINARY; PRT; 553 AA.
 AC 09N0X6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CDNA FL11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10
 DE PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Tsouai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Makamatsu A.,
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
 RT "NDO human cDNA sequencing project."
 RL Submitted (FEBS-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK001933; BAA91986.1;
 DR InterPro; IPR001810; F-box.
 DR Pfam; PF00646; F-box; 1.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 553 AA; 62280 MW; CA829C221986A3F2 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 553;
 Best Local Similarity 30.8%; Pred. No. 1.6e-44;

Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KPNSSEREDCNNGEPPRKIIPEKNSLRQTVNSCARLQNOETVCLASTAMKTENCYAKT 73
 DB 17 KLDHGEVRSFSLAKKPKCV-----SEYTTTGL-----VPCSA-----PTTGGDL 58
 QY 74 KLANGTSMIVPKGRKLSAYEKEL--CYKYPEOMSESDQVEFEHLISQMHYGH 131
 DB 59 RAAGGOG-----GQRRRTSVQPTGLQEWLKMFOSSGGEKLLADELIDSCPTGVKH 113
 QY 132 INSYLKPMLODFETALPARGLDIAENILSYLDAKSICAAELVCKEMRYVTSQGMK 191
 DB 114 MMYVIEPQFQDFISLIP-----KELALYVLSFLEPKDILQAQTCRKWRLLAEDNLLMRE 169
 QY 192 LIEMVWRTSLMGLAF-----RRGWQOYLFKNKPRPDGNAPPNPSFYALYPKIIIDTETI 246
 DB 170 KCKE-----EGIDBPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ--HRI 207
 QY 247 ESNMRGRHSIORIHCHSEFSKGYCCLOYDQKTVSGLRDNTIKIMDKNPLECKRIITGH 306
 DB 208 DTMNRKELASPKV-LKGDHDDVITTCLOFCGNRLVSGSDNITLKVAVAGKCLIFLVGH 266
 QY 307 TGSVLCLOYDERVITITGSDSTVAVMDVNTGEMLNTLIHCEAVLIHLRFNNGMMVYTCSD 366
 DB 267 TGVWSSQMDNIIISGSTDRILKVMNAETGECIHTIYGTSTVRCMHLLEKRVVSGSR 326

QY 367 RSLAVMDASPTDITLRLVIGHRAAVNVDPDKIYVSASGRTIKVMTSTCEVRTL 426
 DB 327 ATLKVMDIEGQCL---HVLGMHAAVRCVQYDGRVRSASAYDEMVKVMDPERETCLHTL 383
 QY 427 NGHRGIACLQYRDRLVYSGSSDNTIRLMDIEGACRLVLEGHEELVRCIRFDNKRIVSG 486
 DB 384 QGHTNRVYSLQFDGIHVHVSGLDTSIRVMDVENGCNHTLTGHSLSGMEKDNILVSG 443
 QY 487 AYDGKIVMDVVALDPRAPAGTLCRTIV---EHSGRVRLQDEFOIVSSSHDDITLI 543
 DB 444 NADSTVAKWIDITG-----OCLQTLQPGPKHOSAVTCLQFNKFEVITSSDDGTIVKL 494
 QY 544 WD 545
 DB 495 WD 496

RESULT 11

QY 044083 PRELIMINARY; PRT; 587 AA.
 AC 044083;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SEL-10.
 GN SEL-10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98051191; PubMed-9389650;
 RA Hubbard E.J.A., Wu G., Kitajewski J., Greenwald I.;
 RT "sel-10, a negative regulator of lin-12 activity in Caenorhabditis
 RT elegans, encodes a member of the CDC4 family of proteins."
 RL Genes Dev. 11:3182-3193(1997).
 DR EMBL; AF020788; AAC47809.1;
 DR InterPro; IPR000412; ABC2_transport.
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00890; ABC2_MEMBRANE; UNKNOWN_1.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 587 AA; 65311 MW; 2D3970B4EFAA1B8C CRC64;

Query Match 17.1%; Score 520; DB 5; Length 587;
 Best Local Similarity 28.8%; Pred. No. 1.4e-34;

Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGTSMI-----VPRKOR--LSASYEKEL-----CYKYPEOMSESDQVE 115
 DB 35 ESSYSNGSSSYNADKLSSRLQHKLDLSASPSRNDLNPVHEHLIALFKLSSABQMD 94
 QY 116 FVEHLISQMHCHOHGINSYLKPMLODFETALPARGLDIAENILSYLDAKSICAAELV 175
 DB 95 APTRLQESNMNTIROLRAIIEPHQDFLCLPV-----ELGMKILHNLGVDLLKVAQV 150
 QY 176 CKRWYVNTSDGMKMLK-LERV-----RDSLMRGLAERKGGQYLFKNKPRPDGNAP 227
 DB 151 SKWKMLITSEDKITWSLGVEEFKNHPDPTDRTVGAMQGTALNAG-----VTIPDHLP 203
 QY 228 PN-SFYRALYPKIIDP-----ETIESNMRGRHSIORIHCRSEFSKGYCYL 273

| | | | | | |
|----|---|-------|------------------|--------|------------------------------------|
| | Query Match | 12.8% | Score 387.5; | DB 11; | Length 304; |
| | Best Local Similarity | | Pred. No. 5e-74; | | |
| | Matches | 84; | Conservative | 72; | Mismatches 119; Indels 27; Gaps 9; |
| OY | 258 QRIHCSEFSSKGYCYLQYD--OKIVSGLRDNITIKIMDKTLECKRLILGHGVSVCIO | 314 | | | |
| Db | 14 ELHETLEGHKNNVYAIAFNPNPYGDRIAGSPFKTKLMASFTGCYHTTFGHAIEVICLS | 73 | | | |
| OY | 315 YDER--VIITGSSDSIVRWVDNTEMTNTLIHCEAVYLHRFNNG--MMVTCSKDRIA | 370 | | | |
| Db | 74 FNPQSIVVAATGMDPTAKLMDIONGEVVTLTGIIAEIISLSPFTSGDRITIGSFHTTV | 133 | | | |
| OY | 371 VPMKASTDTLLRRVLYGHRAVN--VYDFEDKYIVSASGRITKVWNTSTCEVRFLNG | 428 | | | |
| Db | 134 VMDSVGSRKY--HTLIGHCAEISKLRFMDCSLILITSMSKOTCMMDAIRSQRVALITG | 190 | | | |
| OY | 429 HKRCGI--ACLOYRDRLVSVGSSDNTIRLMDIECGACLVLEGHEHYLCIRFD--KKRIY | 464 | | | |
| Db | 191 HDDELIDSCPDYGLKIATAISADGARVYNATRCVYKLEGGHGELSKISFENQGRLL | 250 | | | |
| OY | 485 SGADVGRKYWDVVALMDPRAPGTCLERTIVESHGGRFYLRQPEFO--IVSSSHDPTIL | 542 | | | |

| RESULT | 13 | | | |
|--------|---|--------------|------|---------|
| ID | 096611 | PRELIMINARY; | PRT; | 942 AA. |
| AC | 096611; | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Created) | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Last sequence update) | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | |
| DE | MEK KINASE ALPHA. | | | |
| GN | MRKA. | | | |
| OS | Dictyostelium discoideum (Slime mold). | | | |
| OC | Eukaryota; Mycetozoa; Dictyostellidae; Dictyostelium. | | | |
| OX | NCBI_TaxId=44689; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-KAX-3; | | | |
| RX | MEDLINE=99051319; PubMed=9832508; | | | |
| RA | Chung C.Y., Reddy T.B.K., Zhou K., Firtel R.A.; | | | |
| RT | "A novel, putative MEK kinase controls developmental timing and | | | |
| RT | spatial patterning in Dictyostelium and is regulated by ubiquitin | | | |
| RT | mediated protein degradation."; | | | |
| RL | Genes Dev. 12:3564-3578(1998). | | | |
| CC | -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES. | | | |
| DR | EMBL; AF093689; AAC97114.1; - | | | |
| DR | InterPro: IPR000719; Euk_pkinase. | | | |
| DR | InterPro: IPR001810; F-box. | | | |
| DR | InterPro: IPR002270; OPR. | | | |
| DR | InterPro: IPR002290; Ser_thr_kin_actsite. | | | |
| DR | InterPro: IPR001245; Tyr_kin. | | | |
| DR | InterPro: IPR001680; WD40. | | | |
| DR | Pfam; PF00646; F-box_1. | | | |
| DR | Pfam; PF00069; pkinase; 1. | | | |
| DR | Pfam; PF00400; WD40; 7. | | | |
| DR | PRINTS; PRO0320; GPROTEINBRPT. | | | |
| DR | PRINTS; PRO0109; TYRKINASE. | | | |
| DR | SMART; SM00256; FBOX_1. | | | |
| DR | SMART; SM00016; OPR_1. | | | |
| DR | SMART; SM00220; S_TKC_1. | | | |
| DR | SMART; SM00320; WD40; 7. | | | |
| DR | PROSITE; PSS0181; FBOX_1. | | | |
| DR | PROSITE; PSS00107; PROTEIN_KINASE_ATP_1. | | | |
| DR | PROSITE; PSS0011; PROTEIN_KINASE_DOM_1. | | | |
| DR | PROSITE; PSS0108; PROTEIN_KINASE_SF_1. | | | |
| DR | PROSITE; PSS00678; WD_REPEATS_1; UNKNOWN_4. | | | |
| DR | PROSITE; PSS00082; WD_REPEATS_2; 5. | | | |

DR PROSITE: PS50294; WD_REPEATS_REGION, 1.
 KW ATP-binding; Kinase; Repeat; Serine/threonine-protein kinase;
 KW Transferrase; WD repeat.
 SQ SEQUENCE 942 AA; 105796 MW; C9E4928A8C7C68F7 CRC64;

Query Match 12.0%; Score 365.5; DB 5; Length 942;
 Best Local Similarity 24.7%; Pred. No. 1.5e-21;
 Matches 110; Conservative 67; Mismatches 140; Indels 129; Gaps 16;

QY 175 VCKEYRVTSQGLMKK-----LIERVVRTDSL-WRG-----LAERRGQQLYFKKPPD 223
 DB 547 VCKHMOIIDDDELMMKVCSDRLINKSFESTIMKSNYIKIKQKQW----FHKKL-- 600
 QY 224 GMAPNSFYALKPKIIIDLETIESMRCGRSLQRIHCRSETSKGVYCLQ-YDDQ-KIV 281
 DB 601 -----NHSTLKGH-----DKGVFCVKLIDDOGMVL 625
 QY 282 SGLRNTIKIMD-----KNTLECKR----- 301
 DB 626 SGGEDKKLKWYDISGHHNNHSGIVGSIKSKGLIINNNSNNSSSSSSSSRL 685
 QY 302 -ILTGTSVLCIQ-----DERVIITGSSDSVTRVMDVNTGEMLNTLIHCEAVLHLR 354
 DB 686 FSLKSHSGCIRKSVYQROGSDVSRVFETASADFTCKIFSLTKKTLFTYTNHQAETCIN 745
 QY 355 FNNGM---WVTCSDRSIAVWDMASPTDITLRVLVGHRAAVNVDF-----DKY 402
 DB 746 YLGDVENKCIITSLDKTIQIMDAETGSCLSLTLR--GHGGIYCYKTDVAVTHNGVNH 802
 QY 403 IYSASGDRITKIVWNTSTCEVFTLNGHKGRIACLOVRLVYVSSSDNTIRLMDIECGAC 462
 DB 803 IYASVADKTSNWDTRSSKVSFQHTEDVLCVYFQKVVYTGSCDGIKIMDITGKT 862
 QY 463 LRVLEGHE---ELVRCIFEDKRIKIVSAGYDKIKVMDLVAALDPRAPAGTICLTIVER 518
 DB 863 ISTFLPSETRKNVYVYVDFDSKIISGKTIIRIMDIYNERDSRISIG-----H 913
 QY 519 SGRVRLQDFEFOIYSSSHDDTILM 544
 DB 914 HETIFSLQFNNOKLITGSLDKIKVKTW 939
 RESULT 14
 ID 09PTR5 PRELIMINARY; PRT; 410 AA.
 AC 09PTR5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE L1S1.
 GN L1S1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reiner O., Shmueli O.;
 RT "Characterization of the chicken homolog of L1S1";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113946; AAF18938.1;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPPOTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 410 AA; 46664 MW; FC5848D06EDCA20 CRC64;

Query Match 11.7%; Score 356; DB 13; Length 410;
 Best Local Similarity 29.5%; Pred. No. 3e-21;
 Matches 88; Conservative 56; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLRDNTIKIMDNTLECKRIILGHTGVSVCLOQYDE--RVITGSSDSVTRVMDVNTG 337
 DB 123 IYASASDATIKWYDETGDERTLKGHTDSYODISFDHTGKLKILASCSDMTIKIMDPOGF 182
 QY 338 EMLNLTLIHCEAVLHLR--NNGMAYTCSKRSIAVWDMASPTDITLRVLVGHRAAVNV 395
 DB 183 ECIIRTMHGHNVSSVAIMPNDHIVASARDKTIKMEVQTYCV---KTFGHRVWRM 239
 QY 396 V--DFDDKIYVASGDRITKIVWNTSTCEVFTLNGHKGRIACLOQRD----- 441
 DB 240 VRPNQDGLILASCSDQVIRVAVVATKECKALERHEHNVCEISWAPSSYSTSEATGS 299
 QY 442 -----LVYSSSDNTIRLMDIECGACLRVLEGHEELVRGIRPDN--KRIVSGAYD 489
 DB 300 ETKKSGKRPPELLSGSRDKITIKMWDISTGMCLMTLVGHDMVRGVLFHSGKFLISCAD 359
 QY 490 GKIKYMDLVAALDPRAPAGTICLTIVERHSGRVRLQDFE--QIVSSSHDDTILMD 545
 DB 360 KTLRVWDF-----KNKRCMTLNAHEHFVTSLDFHKTAPYVYVGSVDQTVKWE 408
 RESULT 15
 ID 09GL51 PRELIMINARY; PRT; 410 AA.
 AC 09GL51;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PLATELET-ACTIVATING FACTOR ACETYLAHYDROLASE IB-ALPHA SUBUNIT.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nagasaka T., Bouliday G., Coupel S., Coulon F., Tesson L.,
 RA Heslan J.-M., Souillou J.-P., Charreau B.;
 RT "Cloning of porcine PAF-AH 1b-alpha cDNA and expression in endothelial
 RT cells";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF319658; AAG33867.1;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPPOTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Hydrolase; Repeat; WD repeat.
 SQ SEQUENCE 410 AA; 46654 MW; A08DAFCD88B2719 CRC64;

Query Match 11.6%; Score 351; DB 6; Length 410;
 Best Local Similarity 28.9%; Pred. No. 7.7e-21;
 Matches 86; Conservative 58; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLRDNTIKIMDKTLECKRIILGHTGVSVCLOQYDE--RVITGSSDSVTRVMDVNTG 337
 DB 123 IYASASDATIKWYDETGFERTLKGHTSYODISFDHSGKLILASCSDMTIKIMDPOGF 182
 QY 338 EMLNLTLIHCEAVLHLR--NNGMAYTCSKRSIAVWDMASPTDITLRVLVGHRAAVNV 395
 DB 183 ECIIRTMHGHNVSSVAIMPNDHIVASARDKTIKMEVQTYCV---KTFGHRVWRM 239
 QY 396 V--DFDDKIYVASGDRITKIVWNTSTCEVFTLNGHKGRIACLOQRD----- 441
 DB 240 VRPNQDGLILASCSDQVIRVAVVATKECKALERHEHNVCEISWAPSSYSTSEATGS 299

QY 442 -----LVVSGSSDNTIRLWDIEGACLRVLESGHELVRCIREDN--KRIVSGAYD 489
 Db 300 ETRKSGKPCPFLLSGSRDKTIKMDVSTGCMCLMTLVGHDMWVRGVLPHSGKFILSCADD 359
 QY 490 GKIVMDLVAAALDPRAPAGTILCLRTLVHSGRVFRLOFDEF--QIVSSHDDTILIMD 545
 Db 360 KTLRWMDY-----KNKRCMKTLNNAHEHFVTSIDFHKHTAPYVVTGSVDOTVKWE 408

Search completed: May 8, 2002, 10:53:44
 Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:50:48 ; Search time 13.31 Seconds

(without alignments)
1567.414 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034

Sequence: 1 MDPAEAVLQEKALFKFNSS.....PAAQAEPPSPSRTTYTISR 569

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Database : SwissProt_39:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 3006 | 99.1 | 605 | FWIA_HUMAN | Q9Y297 homo sapien |
| 2 | 2597 | 85.6 | 518 | TRCB_XENLA | Q91854 xenopus lae |
| 3 | 2384.5 | 78.6 | 542 | FWIB_HUMAN | Q9ukb1 homo sapien |
| 4 | 1635.5 | 53.9 | 701 | YSS1_CAEEL | Q09990 caenorhabdi |
| 5 | 690 | 22.7 | 506 | YAF1_SCHPO | Q09855 schizosacch |
| 6 | 590.5 | 19.5 | 605 | POF1_SCHPO | P87053 schizosacch |
| 7 | 575 | 19.0 | 678 | SCOB_EMENT | Q00659 emericella |
| 8 | 545 | 18.0 | 640 | MT30_YEAST | P39014 saccharomyc |
| 9 | 531.5 | 17.5 | 650 | SCO2_NEUCR | Q01277 neurospora |
| 10 | 520 | 17.1 | 579 | SE10_CAEEL | Q93794 caenorhabdi |
| 11 | 477.5 | 15.7 | 684 | CC4_CANAL | P53699 candida alb |
| 12 | 455.5 | 15.0 | 1356 | HER1_PODAN | Q00808 podospira a |
| 13 | 453 | 14.9 | 775 | POPI_SCHPO | P87060 schizosacch |
| 14 | 399 | 13.2 | 779 | CC4_YEAST | P07834 saccharomyc |
| 15 | 396.5 | 13.1 | 703 | POB2_SCHPO | Q14170 schizosacch |
| 16 | 392 | 12.9 | 732 | KMRB_DICDI | P90648 dictyosteli |
| 17 | 354 | 11.7 | 409 | LISI_HUMAN | P43034 homo sapien |
| 18 | 354 | 11.6 | 409 | LISI_MOUSE | P43033 mus musculu |
| 19 | 353 | 11.6 | 409 | LISI_BOVIN | P43033 bos taurus |
| 20 | 341 | 11.2 | 515 | YCM2_YEAST | P25382 saccharomyc |
| 21 | 337 | 11.1 | 1146 | KMRH_DICDI | P42527 dictyosteli |
| 22 | 334.5 | 11.0 | 422 | FBW2_HUMAN | Q9ukt8 homo sapien |
| 23 | 325.5 | 10.7 | 361 | WDS_DROME | Q9v3f8 drosophila |
| 24 | 324.5 | 10.7 | 362 | WDR5_HUMAN | Q9up99 homo sapien |
| 25 | 318 | 10.5 | 422 | FBW2_MOUSE | Q60844 mus musculu |
| 26 | 313.5 | 10.3 | 376 | YKY4_CAEEL | Q17963 caenorhabdi |
| 27 | 312.5 | 10.3 | 714 | YU12_YEAST | P47025 saccharomyc |
| 28 | 307.5 | 10.1 | 742 | PKNA_THECU | P49695 thermomonas |
| 29 | 307.5 | 10.1 | 1693 | Y163_STY3 | Q55663 synecocyst |
| 30 | 306 | 10.1 | 704 | T2D4_DROME | P49846 drosophila |
| 31 | 299.5 | 9.9 | 1194 | APAF_HUMAN | O14727 homo sapien |
| 32 | 298.5 | 9.8 | 800 | T2D4_HUMAN | O15442 homo sapien |
| 33 | 292 | 9.6 | 659 | YK16_YEAST | P36130 saccharomyc |

| | | | | | | |
|----|-------|-----|-----|---|------------|--------------------|
| 34 | 290.5 | 9.6 | 614 | 1 | TU11_SCHPO | Q09715 schizosacch |
| 35 | 287.5 | 9.5 | 327 | 1 | GBLP_BRANA | Q39336 brassica na |
| 36 | 285.5 | 9.4 | 798 | 1 | T2D4_YEAST | P38129 saccharomyc |
| 37 | 283.5 | 9.3 | 377 | 1 | GBLP_ARATH | O24456 arabidopsis |
| 38 | 283 | 9.3 | 473 | 1 | PRP5_SCHPO | O13615 schizosacch |
| 39 | 283 | 9.3 | 713 | 1 | TU11_YEAST | P16649 saccharomyc |
| 40 | 281.5 | 9.3 | 682 | 1 | TU11_KULA | P56094 kluveromyc |
| 41 | 280.5 | 9.2 | 444 | 1 | NODE_EMENT | Q00664 emericella |
| 42 | 280 | 9.2 | 318 | 1 | GBLP_DROME | O18640 drosophila |
| 43 | 277 | 9.1 | 586 | 1 | TU12_SCHPO | Q9uug8 schizosacch |
| 44 | 276.5 | 9.1 | 501 | 1 | YH92_CAEEL | Q23256 caenorhabdi |
| 45 | 273 | 9.0 | 325 | 1 | GBLP_MEDSA | O24076 medicago sa |

ALIGNMENTS

RESULT 1

| ID | FWIA_HUMAN | STANDARD: | PRT: | 605 AA. |
|----|--|-----------|------|---------|
| AC | Q9Y297; Q9Y213; | | | |
| DT | 20-AUG-2001 (Rel. 40, Created) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | |
| DE | F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP) | | | |
| DE | (E3RSIKAPPAB) (PIKAPPALPHA-E3 RECEPTOR SUBUNIT). | | | |
| GN | FBWIA OR FBWIA OR BTRCP OR BTRC. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RX | MEDLINE=99075339; PubMed=9859996; | | | |
| RA | Yaron A., Hatzubai A., Davis M., Layon I., Amit S., Manning A.M., | | | |
| RA | Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.; | | | |
| RT | "Identification of the receptor component of the Ikkappaalpa- | | | |
| RT | ubiquitin ligase."; | | | |
| RL | Nature 396:590-594(1998). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 2). | | | |
| RC | TISSUE=Lymphoid; | | | |
| RX | MEDLINE=98325370; PubMed=9660940; | | | |
| RA | Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V., | | | |
| RA | Thomas D., Strebel K., Benarous R.; | | | |
| RT | "A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu | | | |
| RT | connects CD4 to the ER degradation pathway through an F-box motif."; | | | |
| RL | Mol. Cell 1:565-574(1998). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 2). | | | |
| RX | MEDLINE=20003060; PubMed=10531035; | | | |
| RA | Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M., | | | |
| RA | Pagano M.; | | | |
| RT | "Identification of a family of human F-box proteins."; | | | |
| RL | Curr. Biol. 9:1177-1179(1999). | | | |
| RN | [4] | | | |
| RP | CHARACTERIZATION. | | | |
| RP | MEDLINE=99145464; PubMed=9990852; | | | |
| RA | Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J., | | | |
| RA | Harper J.W.; | | | |
| RT | "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically | | | |
| RT | with phosphorylated destruction motifs in I-kappa-B-alpha and | | | |
| RT | beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."; | | | |
| RL | Genes Dev. 13:270-283(1999). | | | |
| CC | -I- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA | | | |
| CC | (PIKAPPALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR | | | |
| CC | UBIQUITINATION AND DEGRADATION. | | | |
| CC | -I- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX. | | | |
| CC | -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE | | | |
| CC | PRODUCED BY ALTERNATIVE SPLICING. | | | |
| CC | -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN. | | | |
| CC | -I- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS). | | | |

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF101784; AAD08702.1; -
 DR EMBL: Y14153; CAAT4572.1; -
 DR EMBL: AF129530; AAF04464.1; -
 DR MIM: 603482; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
 FT DOMAIN 190 228 F-BOX.
 FT REPEAT 301 338 WD 1.
 FT REPEAT 341 378 WD 2.
 FT REPEAT 381 418 WD 3.
 FT REPEAT 424 461 WD 4.
 FT REPEAT 464 503 WD 5.
 FT REPEAT 505 541 WD 6.
 FT REPEAT 553 590 WD 7.
 FT REPEAT 590 605 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1; Length 605;
 Best Local Similarity 94.0%; Pred. No. 1,1e-219;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAAVILQEKALKFM-----NSSERDC 24
 DB 1 MDPAAVILQEKALKFMCMSPRLMIGCSSLADSMPLSLCLINPGALITATONSSERDC 60
 QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMTENCVAATKLANGTSSMTV 84
 DB 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMTENCVAATKLANGTSSMTV 120
 QY 85 PRORLTSAYEKEKELCYKYPQWSESOVEVEHLISQMKHYOGHINSLKPYLQDF 144
 DB 121 PRORLTSAYEKEKELCYKYPQWSESOVEVEHLISQMKHYOGHINSLKPYLQDF 180
 QY 145 ITALPARGLDHIAENTILSYLDKSLCAAEVLCKEYRYTSGMLMKKLIEMVPTDSLIR 204
 DB 181 ITALPARGLDHIAENTILSYLDKSLCAAEVLCKEYRYTSGMLMKKLIEMVPTDSLIR 240
 QY 205 GLAERRGQGYLFKNKPPDGNAPNSFYRALYPKTIQDIETIESNMWRCGRSLRICHRS 264
 DB 241 GLAERRGQGYLFKNKPPDGNAPNSFYRALYPKTIQDIETIESNMWRCGRSLRICHRS 300
 QY 265 ETSKGVYQLOVDOKIYSGLDNITKIDKNTLECKRLITHTSSVLCLOYDERVIITGS 324
 DB 301 ETSKGVYQLOVDOKIYSGLDNITKIDKNTLECKRLITHTSSVLCLOYDERVIITGS 360
 QY 325 SDSYRVWDVTGEMLNTLIHCEAVLHLRFNNGMAMTCSKDRSIAYWDNASPTDITLRR 384
 DB 361 SDSYRVWDVTGEMLNTLIHCEAVLHLRFNNGMAMTCSKDRSIAYWDNASPTDITLRR 420
 QY 385 VLGVHRAAVNVVDPDDKYIVASASDRTIKWNISTCEFEVRLTLNKHKGIACTQYRDLVV 444
 DB 421 VLGVHRAAVNVVDPDDKYIVASASDRTIKWNISTCEFEVRLTLNKHKGIACTQYRDLVV 480

QY 445 SGSSDNTIRLMDICGACRLVLEBHELYVRCIRFDNRKRIYSGAYDKIKYWDVLAALDPR 504
 DB 481 SGSSDNTIRLMDICGACRLVLEBHELYVRCIRFDNRKRIYSGAYDKIKYWDVLAALDPR 540
 QY 505 APAGTLCRLTVESGVRFLQFDEFOIVSSSHDTLLIMDFLNDPAAQAAPPSPRSRTY 564
 DB 541 APAGTLCRLTVESGVRFLQFDEFOIVSSSHDTLLIMDFLNDPAAQAAPPSPRSRTY 600
 QY 565 TYISR 569
 DB 601 TYISR 605

RESULT 2
 TRCB_XENIA STANDARD; PRT; 518 AA.
 ID TRCB_XENIA
 AC Q91854; P70037; P70038;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
 GN FBXW1 OR BTRCP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9330289; PubMed=8393141;
 RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
 RT anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
 RT with beta-transducin repeats".
 RL Mol. Cell. Biol. 13:4953-4966(1993).
 RN [2]
 RP SEQUENCE OF 302-518 FROM N.A.
 RX MEDLINE=97109804; PubMed=8952061;
 RA Hudson J.W., Alarcon V.B., Elinson R.P.;
 RT "Identification of new localized RNAs in the Xenopus oocyte by
 RT differential display PCR".
 RL Dev. Genet. 19:190-196(1996).
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
 CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
 CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
 CC TADPOLE EMBRYO.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M98268; AAA02810.1; -
 DR EMBL: U63921; AAA49671.1; -
 DR EMBL: U63922; AAA49672.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.

[illegible]

YAF1_SCHPO STANDARD: PRT: 506 AA.

ID YAF1_SCHPO 009855: 09P7V1: 01-FEB-1996 (Rel. 33, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHEICAL 58.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN C2966.01 IN CHROMOSOME 1.

GN SPAC2966.01 OR SPAC30.05.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN NCBI_TaxID=4896;

RP SEQUENCE FROM N.A.

RC STRAIN-972:

RA McDougall R.C., Rajandream M.A., Barrell B.G., Brown S., Murphy L., Jones L., McNeil A., Harris D., to the EMBL/GenBank/DBJ databases.

RL Submitted (JAN-2000).

RN SEQUENCE OF 240-506 FROM N.A.

RA Jones L., Murphy L., McNeil A., Simpson I., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V., Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: A1136538; CAB66464.1; -

DR EMBL: 266525; CA91423.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

DR PRINTS: PR00320; GPROTEINBPT.

DR SMART: SM00256; FBOX; 1.

DR SMART: SM00320; WD40; 7.

DR PROSITE: PS00181; FBOX; 1.

DR PROSITE: PS00678; WD_REPEATS_1; 3.

DR PROSITE: PS50082; WD_REPEATS_2; 6.

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.

KW Hypothetical protein; Repeat; WD repeat.

FT DOMAIN 70 116 F-BOX.

FT REPEAT 219 256 WD 1.

FT REPEAT 259 298 WD 2.

FT REPEAT 301 338 WD 3.

FT REPEAT 345 386 WD 4.

FT REPEAT 426 464 WD 5.

FT REPEAT 427 464 WD 6.

FT REPEAT 468 505 WD 7.

SO SEQUENCE 506 AA; 58257 MW; CER34D4EFBCE10 CRC64;

Query Match 22.7%; Score 690; DB 1; Length 506;

Best local similarity 30.4%; Pred. No. 8.3e-45;

Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 ENCVAR-----TKANGTSSNIIVPKOKKLSAYKEKELCVKFEQMSSEDQVEFVHLIS 122

DB 8 KNVASVSVSDLTSCSDSTSPVCLNPLS-----HENNRIDLRLDLA 50

QY 123 QMCHYHGHTINSYKPMLOGLFTALPARGLDHIAENILSYLDKSLCAAEIYCKEMRYV 182

DB 51 SLKREGVAVNIVNRSLFTFDTEVFP-----EVSILRVSYLDQDLCKCKLMSKRMKRL 106

QY 183 TSDGMLMKKLI-----ERMVFTDSLWRG-----LAERRGNG----- 213

DB 107 LEDGIMKALYMKQGWFEVNEVLNEFEAMRTRHKFPQPRENFKKOONIGPYGMFLQ 166

QY 214 QYLEKKNPPGQANPPNFFYALPKIQTIDLETIESMWRCSRHLQRHCS----- 264

DB 167 QEIF-----DSNGRPLNWSLY--KEHAHLDNSNRHGRFLVSTFNPNSTIRPADQ 217

QY 265 -ETSKGVCLQYDQKIVSGLRDNTIKIMDKNTLECKRIILTGHVGLQYDER--VII 321

DB 218 RATIDSVYCYQYDEIVSGSKDRTVSMDVNSFLLKLYHSGSVLCIDFCRRRLV 277

QY 322 TGSSDSYVRWVDVTGEMLNTLHCEAVLHLRFNNGMAYTCKDSIAYV--DMSPTD 379

DB 278 SGSSDSITIIIDMQRRLPKYFGHTDNVGVVSENYIISSSDRTRARWRDATSPAE 337

QY 380 ITRRLVYGHRAAVNVYDEDDK--YIVASGDRIRKWNSTCEFTVTLNGHKGIACQ 437

DB 338 ACM-HVLRGHLASVNSVOYSKGTGLIVASDRKLRTWDLTTGHCITIIAHORGICAQ 396

QY 438 YRDLRVSGSSDNTIRLMDIEGACLVLEGHELVRCIRDNKRIYSGAYDKIKYMDL 497

DB 397 YNGKFIVSGSSDLIRIFEASSGKLRMLQGHEDLIRTVRFNDEKIVSGYDGTVRIMN 455

QY 498 VVALDPRAPACTLCRLIVER-----SGRVFLQDFEQIVSSSHDPIILWDF 546

DB 456 -----FNTGECHVLHNSRNSRVFGLQDFHRIITACTHSSEILWNF 497

RESULT 6

POFI_SCHPO STANDARD: PRT: 605 AA.

ID POFI_SCHPO P87053; 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE F-BOX/WD-REPEAT PROTEIN POFI.

GN POFI OR SPAC57A10.05.C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN NCBI_TaxID=4896;

RP SEQUENCE FROM N.A.

RC STRAIN-972:

RA Katayama S., Kitamura K., Toda T.; Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

FT fission yeast.

FT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.; Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: AB032410; BA84528.1; -

DR EMBL: Z94864; CAB08168.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 2.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 107 153 F-BOX.
 FT REPEAT 271 299 WD 1.
 FT REPEAT 311 339 WD 2.
 FT REPEAT 350 379 WD 3.
 FT REPEAT 390 420 WD 4.
 FT REPEAT 432 460 WD 5.
 FT REPEAT 472 500 WD 6.
 FT REPEAT 510 538 WD 7.
 SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;
 Best Local Similarity 29.0%; Pred. No. 3,5e-37;
 Matches 156; Conservative 86; Mismatches 169; Indels 107; Gaps 12;

QY 80 SSNIYPKQKSLASYTEKEKELCVKFEQWSE---SDVEVEHLISOMCHYQHCHINSYL 136
 DB 46 SSM---HNELSGLEKSKORVAVAAFAFSEASCERKIALQGLINNCSSLSLSPASTL 101
 QY 137 KPMLODFITLPAAGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMKKLIERM 196
 DB 102 DSLVLDLFLSLPV---EISFRILSFIDASLSCAAQVSKHMKFLADDVYIMRMCQH 157
 QY 197 VATDSLNRGLAERRGMOYLKFN----- 219
 DB 158 INRK-----CEKCGMGLPILERRNTLYAAKASIQKRYERLTKRGVDAQHESSPYKAKLD 211
 QY 220 -----KPPGNAPPNSE-----YRALPKIIIDLETIESMWRGRHSLO 258
 DB 212 DYPITSSNETTSSVAPSPNSDSKFEPLPKTRPKVEVAERCR---VECNMRGR--- 263
 QY 259 RINCR---SETSKGVYCLQYDDQKIYSGRLDNTIKIKDKTLCKRILTGHTSVCLQ 314
 DB 264 ---CROVVLGSHDGMCLQVLRNINLASGYDATIRLMNLTPOVALLESHSGSVCLQ 320
 QY 315 YDERIITIGSSSTYRVWMDVTGEMLNTLIHCEAVLIHLRNNGMVTCSDRSIAVDM 374
 DB 321 FDCKLIGSGMDKIRIRNNYRTSEICISILHGTDSVLCITPDSLTVSGSADCTVAKLHF 380
 QY 375 ASPDITLRVLYGHRVAVNVDF--DDKYTVASGDRITVWMTSCPEFRTLNGHKG 432
 DB 381 SSGKRITLR---GHTGVNSVRIIRDRGLVLSGSDSTIKIWSLENTCLHTSAHGP 436
 QY 433 IACLOYRDLRVVSGSSDNTIRLMDIEGACLVLEGHEELVRCIRFDRKRVSGAYDGI 492
 DB 437 VQSLALADSRFLFSCSLDTIKOMDIEKKKCVHTLFEGHGEWETJADHLRLISAHGAV 496
 QY 493 KVMVDVAALDPRAGTICLTILVEHSGRVRLQDFEQIYSSSHDITILMDLNP 550
 DB 497 KVMEECE-----CVHTLKNHSEPTVALGDEEVSGSDCKITVLPNNAP 543

RESULT 7
 SCOB_EMENT
 ID SCOB_EMENT STANDARD; PRT; 678 AA.
 AC 000659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SULFUR METABOLITE REPRESSION CONTROL PROTEIN.
 GN SCOB OR MAP1.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiata; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;

FN [1]
 RE SEQUENCE FROM N.A.
 RA Notoif R.
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
 CC REPRESSION.
 CC -1- SIMILARITY: CONTAINS 8 WD_REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC EMBL: U01220; AAC15905.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR SMART; SM00256; FBOX; 1.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Transcription regulation; Repeat; WD repeat.
 FT DOMAIN 178 224 F-BOX.
 FT REPEAT 347 375 WD 1.
 FT REPEAT 387 415 WD 2.
 FT REPEAT 427 455 WD 3.
 FT REPEAT 466 496 WD 4.
 FT REPEAT 508 543 WD 5.
 FT REPEAT 553 595 WD 6.
 FT REPEAT 607 635 WD 7.
 FT REPEAT 647 675 WD 8.
 SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;

Query Match 19.0%; Score 575; DB 1; Length 678;
 Best Local Similarity 26.9%; Pred. No. 6e-36;
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

QY 117 VEHLISOMCHYQHCHINSYKPMLODFITLPAAGLDHIAENILSYDAKSLCAELVC 176
 DB 153 LOGIMAQCCFPOLSYISATVRLIRIDFTALP---PEIAFKIICLYDITSLCKASQVS 208
 QY 177 KEMRYTSGMLMKKLIEMVVRTDSLMRGLAERRGMOYLKFN----- 220
 DB 209 KGRALADDDVYVHNRKCEQHHRK-----CKKCGMLPLDRKRLRSKRETELRAVTV 262
 QY 221 -----PPDGA--PP-----NSFYRLY---P 237
 DB 263 DRGVVPRSPDASAEPSGKRKLEDDVAVYKRRCSLSGSDAGVDKSDFFKTRYRPMK 322
 QY 238 KIIDLETIESMWRGRHSLOIRHCEVSK-----GYVCLQYDDQKIYSGRLDNTIKIWD 293
 DB 323 EYIKDFKVGITMKYGR-----CSIKFKGHTNGVLCQEDNIIATGSDYTIKTIWD 375
 QY 294 KNTLECKRILTGHTSVCLQYDERIITIGSSSTYRVWMDVTGEMLNTLIHCEAVLHL 353
 DB 376 TENGELFRLRHSIESIRLOPDKILISGMDRIKVMNRTGSCITTYGHGAVIGL 435
 QY 354 RNNGMVTCSDRSIAVMDASPTDITLRVLYGHRVAVNV--VDFDKYIVSAGDRT 411
 DB 436 HFDASILASGVADKTVKINFEKSTFSLR---GHTDMVAVNAVVDYSSRTVFASDDCT 491
 QY 412 IKVWMTSCPEFRTLNGH-----KRGIA----- 435

```

DB 492 VRLMDLDTKTCIRTFHGHVGOVQVPLPRFEFEHHAECENDLSTTSGDANPPSIOA 551
      ::::: | :| :| :|
      436 -----10YRD-----RLVYSSGSSDNTRIRLMDIEGCACLVRLGHEEL 472
      DB 552 SNGLEPNAVSSSASAFGSPDNGRAAPPYRWTSALDSTIRLMTETTGTCILTFPGHLEB 611
      473 VRCIRDPNRIYSGAVDGRKIKYMDLVLAALDPPRPAAGTLCRLTVHSGRVRFLQDFEQI 532
      DB 612 VVALGADTRIRIYSGAEDNRKIKIMD-----PRTGKERFETGSHGPGVTCIGLDSRF 662
      QY 533 VSSSHDDTLLIMDF 546
      DB 663 ATGSEDCVEMRYSF 676

RESULT 8
M30_YEAST
ID M30_YEAST STANDARD; PRT; 640 AA.
AC P39014;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE MET30 PROTEIN
GN MET30 OR YIL046W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=42180-1A;
RA MEDLINE=96069360; PubMed=8524217;
RA Thomas D., Kuras L., Barbey R., Charest H., Blaiseau P.L.,
RA Surdin-Kerjan Y.;
RT "Met30p, a yeast transcriptional inhibitor that responds to S-
RT adenosylmethionine, is an essential protein with WD40 repeats."
RL Mol. Cell. Biol. 15:6526-6534(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrill B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajadream M.A., Riles L., Rowley N., Skellern J., Smith V.,
RA Welsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
CC GENES EXPRESSION.
CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z46861; CAAB6905.1; -
CC EMBL; L26505; AAA96717.1; -
CC SGD; S0001308; MET30.
CC InterPro; IPR001810; F-box.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00646; F-box; 1.
CC Pfam; PF00400; WD40; 6.
CC PRINTS; PR00320; GPROTEINBRPT.
CC SMART; SM00256; FBOX; 1.

```

```

DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Methionine biosynthesis;
KW Cysteine biosynthesis; Repeat; WD repeat.
FT DOMAIN
FT 181 227
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT REPEAT 61 61 M -> I (IN REF. 1).
FT CONFLICT 61
SQ SEQUENCE 640 AA; 72835 MW; 5135D4BCA2E1EB97 CRC64;

Query Match 18.08; Score 545; DB 1; Length 640;
Best Local Similarity 28.98; Pred. No. 1e-33;
Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKSLRQTYNSCARLCLN-QETVCLASTAMKTENCYAK-TRLANGTSMIVPKORKLS 91
DB 86 LPEVFTKFCYRHNPDIPSPPTNACRYKODLKRQELINANAKLPLOQSDIHIIISKYS 145
QY 92 ASYEKEKEICVAYFEQWSESDQVEFEHLISOMCHYOHGINSYLYKMLQDFITALPAR 151
DB 146 NSNDKIRKL-----ILDGLISTSCFPOLSTYSSLVTHMIKIDFISILP-- 188
QY 152 GLDHAENILSLDAKSLCAAEINCKEYRYTSPGMIMKLIEMVRRDLSMRGLAERR- 210
DB 189 --QELSLKILSLDQSCNATVCRKWKQKLADDRVYHACEOH-----DRKC 236
QY 211 ---GMGYLFFKKN-----PPDGNAPNSFYKALPKIYDIETIESNMRCGRSLQ 258
DB 237 PRCGNGPLLHKRRARIRIQONSTGSSNMADIQTQTRPKVYIRREFKVESNMRRG---- 291
QY 259 RIHCSESESK---GVCLQYDDQKIVSGLDNDTIKIMDKNTLECKRLITGHTSVLCLO 314
DB 292 --HCRIOEFGKGMDDVLLLOEFNYRLFLFTGSYDSTIGIMDLFTGKLIRLSGHSOGVKLY 349
QY 315 VDERVITIGSSSYRVMDVNGEMNLNLHHCSEVVLHFRNNGCMVYCSKDRSIAYWDM 374
DB 350 FDDRKLITGSLDKTTRVMWYITIGECISTYRGHSVLSVDSYQKVIYSGSADKIVKVMHV 409
QY 375 ASPTITLRLVLRVGHRAAVNVVDFDKYI--VSASGDRITIKWNTSGCFYRTINGH-- 429
DB 410 ESRCTYTLR-----GHEWVNCVKLHPKSPSCSCSDDTTIRMMDIRITNSCLKVFRGHVGO 465
QY 430 KRGIACLOYRD--RLVSSGSS----- 448
DB 466 VOKIIPITIKIDVENLATNTSDSSSPDDPTWTGDADESDPPSNQETVLDENITPYPTHL 525
QY 449 -----DNTIRLMDICGACRLVLEGHHEELRIRCDNRRIYSGAVDGRKIKYMDLVLAALDP 503
DB 526 LSCGLDNTIKLMDVTKGKIRTFQGHVBEWDIADNRIISGSHSDGSIKYMWDLQSG--- 582
QY 504 RAPAGTLCRLTVHSGRVRFLQDFEQIVSSSHDDTI 541
DB 583 -----KCMHTF---NGR--RLQRETHQHTQPSLQDKV 609

RESULT 9
SC02_NEUCR
ID SC02_NEUCR STANDARD; PRT; 650 AA.
AC 001277;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE SULFUR CONTROLLER-2 (SCON2).

```

```

GN SC0N-2.
OC Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxId=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241499; PubMed=7724564;
RA Kumar A.; Paletta J.V.;
RT "The sulfur controller-2 negative regulatory gene of Neurospora
crassa encodes a protein with beta-transducin repeats.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
CC -!- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -!- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MET30/SC0N-2 FAMILY OF WD-REPEAT
PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17251; AAA68968.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 1.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repeat; WD repeat.
FT DOMAIN 124 170
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

```

Query Match 17.5%; Score 531.5; DB 1; Length 650;
 Best Local Similarity 25.0%; Pred. No. 1,1e-32;
 Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

```

QY 117 VEHLISOMCHYQHGHINSYKPKLQRFDTALPARGLDHTAENITSLYDAKSLCAAEIYVC 176
DB 99 LQGLISOLCPPLSPFVSRVENALKIDFISALPV---ELAQVYLCLDITVSLTKAAQVS 154
QY 177 KEWRYVTSDCMLMKLEKRVNRDLSLMRGLAEKRGQYLFKKRP-----P 222
DB 155 QRRMTLADSAVAVYRMCEQHVNRK-----CTKCGMGLPFLERKLLNRYTRQRLAKGGP 208
QY 223 DQNA-----PNSFYRA 234
DB 209 QGRVTELDASHDSQDSVNOHGKRPAAEAEDPPIKKRQMAAASKAVENTQKTSWMA 268
QY 235 LYKRIIDITISNMRGSHSIQRHCHSETSKGYCYQYDDOKIVSGIRNDTITKIMDK 294
DB 269 VY-----RDRQVSYNKNKSHRYKLTVL---KGHNGVTCLQDLDNIIATGSDYPTTIKIWI 321

```

```

QY 295 NTLCEKRLILGHTGSVLCLOYDERVITGSSDSTVRVMDVNTGEMLNTLIIHCEAVLHR 354
DB 322 ETEECITTLWGHAGIRALQFDDSKISGLDHTIKYWMHTGEICLSTFAAHDVSIVH 381
QY 355 FNNGMWVTSKDRSIAVMDASPTDILRVLVGHAAVNV--VDDDDKIYISASCDRTI 412
DB 382 FDGHLLASGSSDKTVKIFDENSEKTYCLK---GHSQWVNSTHVDIKSRVFSASDDTTI 437
QY 413 KVMNTSTCEVFRTKLNGHKRGI-----ACLQYRDLRVVS----- 445
DB 438 KLMDLDRQYIRFYEYGHVHVOVLILPREYEPDEVINGASQDQNDAMSVSSGSGSPS 497
QY 446 -----GSSD 449
DB 498 MSHAQIERAGSPGSHSSSHNLPPSSLPSCDEDEVRLHLYGSAFAVADSRPLPPRYFMTGGLD 557
QY 450 NTRLMDEIACGLRYLEGHLEELVRCIRPDNKRIVGAYDGKIKVMDLVAALDPAPAGT 509
DB 558 STMRLMDSATGRCLRTFLFGLBEGVWSLADDTIRVIGANDGVKTYE-----PRSG 608
QY 510 LCRLTVHSGRVRLQFDEPQIVSSHDDTILIMDF 546
DB 609 KCDATYTGCGPVCVGLSDSIMASGSEDGTTIRLSHF 645

```

RESULT 10
 SEQ1_CAEEL
 ID SEQ1_CAEEL STANDARD; PRT; 579 AA.
 AC 093794;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SEL-10 PROTEIN
 GN SEL-10 OR F55B12.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Sims M.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RC STRAIN-BRISTOL N2;
 RA Jones S.J.M.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z79757; CAB02129.1;
 DR WormRep: F55B12.3; CE16120.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 5.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.


```

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: Y08391; CAA69671.1;
CC EMBL: AL022103; CAA17898.1;
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00646; F-box; 1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PRO0320; GPROTEINBRPT.
CC SMART: SM00256; FBOX; 1.
CC SMART: SM00320; WD40; 6.
CC PROSITE: PS00181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 3.
CC PROSITE: PS00082; WD_REPEATS_2; 4.
CC PROSITE: PS0294; WD_REPEATS_REGION; 1.
CC Repeat; WD repeat.
CC KW DOMAIN 298 345 F-BOX.
CC FT REPEAT 444 472 WD 1.
CC FT REPEAT 484 538 WD 2.
CC FT REPEAT 575 603 WD 3.
CC FT REPEAT 615 645 WD 4.
CC FT REPEAT 657 687 WD 5.
CC SEQUENCE 775 AA; 87816 MW; B06EDBA46553EEC1 CRC64;

Query Match 14.9%; Score 453; DB 1; Length 775;
Best Local Similarity 25.8%; Pred. No. 1.2e-26;
Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

QY 9 QEALFEMNSEREDC-----NNGEPKRIPEKNSLRQTY-NSCARLCLNQET 56
DB 121 REKCLRRNSLSNLHANKRFLFNSGDSGNKKKEFPSTNINNVYPPNCKSEKVASSE 180
QY 57 VC-----LASTAMKTENCVAATKLANGSSMTYVKORKLSA----- 92
DB 181 TFSLDAPNNSVNSYSPNLLGDSKTRQSPPHSSSSSHNSLHEVIYDFSENPISHP 240
QY 93 -----SYEKREL-----CYKFEQMSDSDOYEFVHILISOMCHQGHINSYKLPMLQROF 144
DB 241 SNHLSQKNAVLKLAQLISFEKLPESVROYLLPHLSRCGKAHAYONIRKILLPFOKNF 300
QY 145 ITALPARGDLHIAENILSYLDASLCAAEMLCKEWTYR-VTSDGMLMKT-----I 193
DB 301 LTGPDA-----ETNLYVTHIDAPSLCAVSGVSHHWYKLVSSNELMKSFLDKGPFWDI 356
QY 194 ERWVRDLSMRGLAERRGWQYLFKKRPPGNAFPNSFYALYPKIIOLETIESMWRG 253
DB 357 DSKIRFMCLRQSLSA-----CAIMKRYFERHFNLERM--- 389
QY 254 RHLQRIHCHSESKGYVC---LOYDOKIVSLRNTIKIMDKNLECKRIILGHTGSV 310
DB 390 LHAPEKIKRCSFPIHGRLLTKQFDDDKITIVTSCSPRINIYTKGVILRSLEHGGDV 449
QY 311 LCLQYDERVITGSSSDSTVAVMDVNTGEMLTLIHCEAVLHLPFNGMVAWVCSKDRSIA 370
DB 450 WTEFYVDTLVGSTRVAVMDLRTECKQVYGHSTIRCIKIYVOGNOSTDTD----- 505
QY 371 VMDASPTDITLRVAVGHAANVVDKXYVASGDTITVWN----- 416
DB 506 -----DVEKE-----NRPASDANSMPYIITSSRDCYTRIMSLPCLDPPVAVNE 552
QY 417 -----TSTCEPVLTLNCHK---RGIAQLQYDRDLVAVSGSSDNITRLMDIEGAC 462
DB 553 NPQDNDFTSATITPFIYIRLRLRGHTDSYREVACL---GDLIVASDYDGTILRVAKASGVC 609

```

```

QY 463 LRVLEGHELVRCIRPDNKR--IVSGAYDGKIKYWDLVVALDPAPAGTCLRTLVHSG 520
DB 610 LHLVLRGHGRVYSTVITNPSRQOCISAGDAKIRIMNL-----ESGEL-LQTLHGSHN 660
QY 521 RVFRLQFDEPQIVSSS--HDDTILIMD 545
DB 661 LVSQVTFNONTIVASAPDPTSLRWMD 687

RESULT 14
CC4_YEAST STANDARD; PRT; 779 AA.
ID CC4_YEAST
AC P07834;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 4.
GN CDC4 OR YFL009W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88011240; PubMed=3309335;
RA Yochem J., Byers B.;
RT "Structural comparison of the yeast cell division cycle gene CDC4 and
RL J. Mol. Biol. 195:233-245(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae."
RN [3]
RP Net. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Churcher C., Rajandream M.A.;
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOKULATION. REQUIRED FOR HTAI-HTBI LOCUS
CC TRANSCRIPTION ACTIVATION.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: X05625; CAA29113.1;
CC EMBL: D50617; BAA09229.1;
CC EMBL: D31600; BAA06495.1;
CC EMBL: Z46255; CAA86341.1;
CC PIR: A26867; A26867.
CC SGD: S0001885; CDC4.
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00646; F-box; 1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PRO0320; GPROTEINBRPT.

```

DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS50082; WD_REPEATS_2; 5.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
 FT DOMAIN 272 319 F-BOX.
 FT REPEAT 380 408 WD 1.
 FT REPEAT 420 449 WD 2.
 FT REPEAT 461 493 WD 3.
 FT REPEAT 528 556 WD 4.
 FT REPEAT 568 598 WD 5.
 FT REPEAT 630 658 WD 6.
 FT REPEAT 669 698 WD 7.
 FT CONFLICT 460 460 K -> E (IN REF. 1).
 SQ SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CAC64;

Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 1.4e-22;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

OY 56 TWCLASTAKTENCYAKTFLANGSTSMIVPKRKLSASYEK-----ELCYVFQW 108
 DB 194 TPLAKTYTINN-----NNIADLESKDSITISPELSEIFSAINNMPHAYFK-- 244

OY 109 SESDOVEFVHLISOMCHYHGHINSYKPMLODFITALPARGDHAENILSYLAKS 168
 DB 245 -----NLLFRLVAMNDSSELDGTLIKDNKRLITSLP-----ELSKIFNLOED 294

OY 169 LCAAEIVCEKWR-VTSDGMLKKLI--ERWRTDLSLWGLAERRGWQYLFFKNKPPDGN 225
 DB 295 IINSVSQNMWKIKRSTSLMKLLISENFV-----SPGKF 331

OY 226 APPNSFYRLVYKIID-----IET--ISNMRGRHSLSQIRHCRSETSKGYCLOYD 276
 DB 333 NSLNKLKSOCTPKLSQOORLISFLENFLLKNMYPNPFVQRTTLRGHMSVITCLOFE 391

OY 277 DQIVSGLRDNTIKIMDKNTECKRILGHGVSYLQYDE-RVLIITGSSDSTVRWMDVN 335
 DB 392 DNYVITGADDKMIRYDSINKKFLQLQSLGHGQWALKYANGILVSGSTRTYRWMDIK 451

OY 336 TGEMLNTLIHCEAVLHLRFNNGMAYVTSKDRSLAVWDMASPTDITLRLVLVGHRAAVN 395
 DB 452 KG-----CCT-----HYFKGHNSTVRC 468

OY 396 VFDD-----KIYVASGDRITKVMNT-----STCE-----FVRLING 428
 DB 469 LDIVYKRIKITYGSSRNTLHMKLPRESSVPPDHGEHDYPLVPHPEENPYFVGVLRG 528

OY 429 HKRGIALOYRDLVVGSSDNTIRLMDIEGACLRVLEGEHELVRCIFEDN--KRIYSG 486
 DB 529 HMAVRYVSGHGNIVSSGYDNTLIVMDVQMKCYILSGTDLISYIVYHEKRCISA 588

OY 487 AYDGKIKYWDL-----VAALDPRAPATL--CLRTLVHSGRFRLODFEQVYSS 535
 DB 569 SMDTIRIWDENIMWNGECYATNSAPCAKILGAMVYLOGHTALVOLRLSDKPLVSA 648

OY 536 SHDDTILIMDFLPPAAQAEPSPRSRYTY 566
 DB 649 AADGSIRGMD-AND-----YSKRFST 668

RESULT 15
 POP2_SCHPO STANDARD; PRT; 703 AA.
 AC 014170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE WD-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).
 GN * POP2 OR SUD1 OR SPAC4D7.03.

OS Schizosaccharomyces pombe (fission yeast)
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes;
 OX NCBI_TaxID=4896;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wolf D.A., Jackson P.K.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RP 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "audit targets cyclin-dependent kinase-phosphorylated Cdc18 and Rnm1
 RT proteins for degradation and stops unwanted diploidization in fission
 RT yeast."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN 13
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: PUTATIVE CYCLIN-DEPENDENT OF E3 (PROTEIN-UBIQUITIN LIGASE)
 CC COMPLEX: TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND
 CC RUM1 PROTEINS FOR DEGRADATION.
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC - SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF038867; AAC95480.1; -
 DR EMBL; AF064515; AAC39496.1; -
 DR EMBL; Z98602; CAB11275.1; -
 DR InterPro; IP001810; F-box.
 DR InterPro; IP001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS50082; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 236 283 F-BOX.
 FT REPEAT 368 417 WD 1.
 FT REPEAT 429 473 WD 2.
 FT REPEAT 505 533 WD 3.
 FT REPEAT 545 575 WD 4.
 FT REPEAT 587 615 WD 5.
 FT REPEAT 625 654 WD 6.
 SQ SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CAC64;

Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. No. 1.9e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

OY 14 KPNNSERED--CNNGEPRKI-----IPEKSLKQTYSCARLCL----NOETVCL 59
 DB 96 KPNESIENVSYKCLDHPDSVPGDFSIISIVPQRNFL-YSHSLPCKIISIDRNRIKID 154

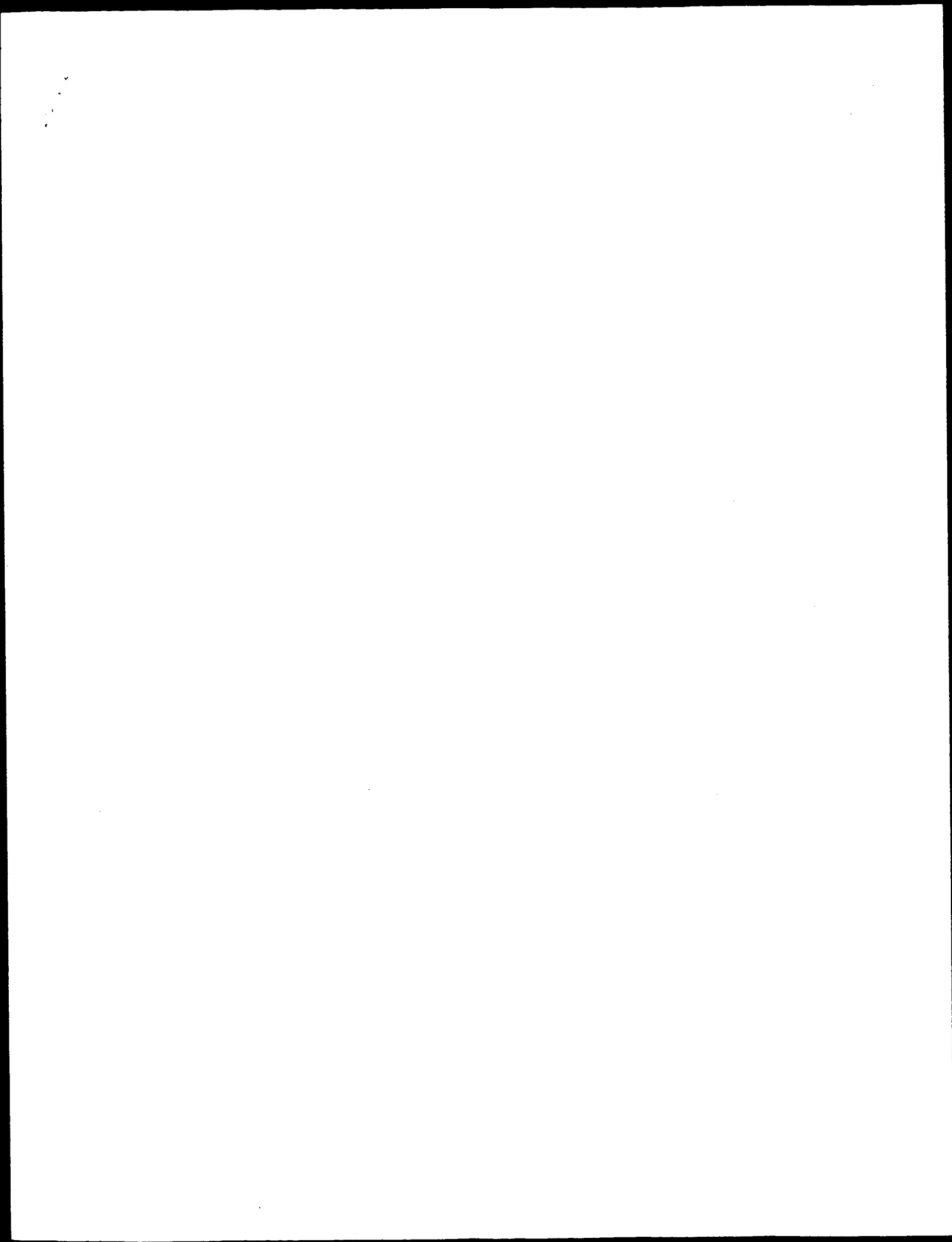
OY 60 ASTAMTKENCYAKTFLANGSTSMIVPKRKLSASYEK-EKELCYVFQWGESDQVEVE 118
 DB 155 NSIISNSDNFPPSPKV--DTSNTVSPGSKPISEDLIDNLQSIYOTFEDLPBGIO-STAF 211

```

QY 119 HLISOMCHYQGH--INSYLKPMQLORDITLALPANGLDIHENILSYLDAKSLCAAEYCK 177
Db 212 FOLMSCRQSMRMLLINCCEPLLKKDILSNLP-----SIVOSTILNLDIHSLFCLSLVSP 267
QY 178 EMYRY-----TSDGMILMKLLEBMRYRTDLSIRGLAERGMQGYLFKKRPDGNAPP-----N 229
Db 268 TWNRRLDVHTS--YKHKMF-----SLGFOQINENDM-----KYPANPLNRPPLFLND 312
QY 230 SFYRALYPKIIIDITETTESMWRGSHLSOR----- 259
Db 313 QISDDPEPELIFK-----RHLNKKRWMLPSPSPHLSFPLIHPNEMITSLLLH 360
QY 260 -----IHCRETSK--GVYCLQYDQCKIVSGLDNTIKIDKXNT 296
Db 361 KDRIIITSGSCTIQIHNAITGVLEARLEGHKEGVMARKHENTLVSGSIDKTVVRNIEK 420
QY 297 LECKRILLTGHTGSVLCLOY-----DERVLIITSSDSTYKRVMDVNTGEML 340
Db 421 AKCHHIFRGHISITRICEILVPSRLRHGVEIYEPQOPYVSGSRDHTLRV----- 472
QY 341 NTLIHHCVALHLFEFNNGMMVATCSKORSLAVMDMASPTDILRLVLYGHRAAVNYVDFD 400
Db 473 -----KLPKNTDPRLPDNTNINSIDRWE-KMRYFP-----HTLIGHTDSKRITSGY 517
QY 401 KYIYASAGDRTIKYWNSTSTCEFWYRTINGHKKRGITACLOYRD--RLVYSGSDNTIRLMDIE 458
Db 518 DIIVSGSYDSSIRIMRYSTEGCELYHIRGHSRLRYSLVLEPERNICISGSMDSIRWMDLS 577
QY 459 CGACRLVLEGEHEELRYCIRFDNKRRIYASGADGKIKWMDLYALMDPAAPAGTCLRLTYEH 518
Db 578 TGTCIKYVLEBGHDAFYTLTIANFQNRLLISGASDSTIRLMD-----LNTGKP-----LAVLPDN 628
QY 519 SGRYFRLLQDFEFOIVSSGSHDTTLLMD 545
Db 629 SGYISSFVSDEHKII--SGNDGYSVKLMD 654

```

Search completed: May 8, 2002, 10:54:21
Job time: 213 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 8, 2002, 10:49:08 ; Search time 19.54 Seconds
(without alignments)
2218.184 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034
Sequence: 1 MDPAAVLOEKALKFNSSSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2597 | 85.6 | 518 | 2 B48088 | beta-transducin re |
| 2 | 1635.5 | 53.9 | 701 | 2 T16607 | hypothetical prote |
| 3 | 690 | 22.7 | 506 | 2 T50211 | WD-repeat protein |
| 4 | 590.5 | 19.5 | 605 | 2 T38832 | probable sulfur me |
| 5 | 545 | 18.0 | 650 | 2 S49932 | MET30 protein - ye |
| 6 | 531.5 | 17.5 | 650 | 2 T46660 | sulfur controller- |
| 7 | 520 | 17.1 | 579 | 2 T22703 | hypothetical prote |
| 8 | 519.5 | 17.1 | 267 | 2 S62507 | hypothetical trp-a |
| 9 | 455.5 | 15.0 | 1356 | 2 T18521 | beta transducin-11 |
| 10 | 453 | 14.9 | 775 | 2 T45136 | WD repeat protein |
| 11 | 399 | 13.2 | 779 | 2 S56245 | cell division cont |
| 12 | 396.5 | 13.1 | 703 | 2 T43557 | F-box/WD-repeat pr |
| 13 | 354 | 11.7 | 409 | 2 S36113 | Lis-1 protein - hu |
| 14 | 354 | 11.7 | 410 | 2 S48052 | platelet-activatin |
| 15 | 341 | 11.2 | 515 | 2 S19487 | hypothetical prote |
| 16 | 339.5 | 11.2 | 777 | 2 T41075 | hypothetical WD-re |
| 17 | 337 | 11.1 | 1146 | 2 A55532 | myosin-heavy chain |
| 18 | 332.5 | 11.0 | 317 | 2 T46032 | WD-40 repeat regul |
| 19 | 330 | 10.9 | 333 | 2 G85034 | probable WD-repeat |
| 20 | 318 | 10.5 | 422 | 2 A56640 | CDC4 repeat unit-c |
| 21 | 313.5 | 10.3 | 376 | 2 T19266 | hypothetical prote |
| 22 | 312.5 | 10.3 | 714 | 2 S56893 | hypothetical prote |
| 23 | 312 | 10.3 | 876 | 2 T51507 | WD40-repeat protei |
| 24 | 307.5 | 10.1 | 1693 | 2 S76086 | beta transducin-11 |
| 25 | 305 | 10.1 | 704 | 2 S33263 | transcription init |
| 26 | 299.5 | 9.9 | 1194 | 2 T03818 | apoptotic proteina |
| 27 | 297 | 9.8 | 659 | 2 T41148 | trp-asp repeat con |
| 28 | 292 | 9.6 | 659 | 2 S38108 | hypothetical prote |
| 29 | 290.5 | 9.6 | 614 | 2 S58306 | WD-40 repeat regul |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 287.5 | 9.5 | 327 | 2 S48839 | guanine nucleotide |
| 31 | 285.5 | 9.4 | 798 | 2 S34023 | TATA box-binding p |
| 32 | 283 | 9.3 | 713 | 2 JN0133 | WD-40 repeat regul |
| 33 | 280 | 9.2 | 494 | 2 T19550 | hypothetical prote |
| 34 | 278.5 | 9.2 | 380 | 2 T40283 | beta-transducin - |
| 35 | 277 | 9.1 | 586 | 2 T38992 | WD-40 repeat regul |
| 36 | 276.5 | 9.1 | 501 | 2 T27513 | hypothetical prote |
| 37 | 274 | 9.0 | 2241 | 2 T16064 | hypothetical prote |
| 38 | 273 | 9.0 | 325 | 2 T09613 | probable GTP-bind |
| 39 | 272.5 | 9.0 | 1008 | 2 T32986 | hypothetical prote |
| 40 | 272.5 | 9.0 | 1019 | 3 JC7538 | neuronal different |
| 41 | 271.5 | 8.9 | 817 | 2 S51445 | probable membrane |
| 42 | 270 | 8.9 | 451 | 2 S65162 | hypothetical prote |
| 43 | 270 | 8.9 | 473 | 2 T33805 | hypothetical prote |
| 44 | 270 | 8.9 | 504 | 2 T50983 | probable pleiotrop |
| 45 | 270 | 8.9 | 651 | 2 T50289 | WD repeat protein |

ALIGNMENTS

RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trcp

C:Species: Xenopus laevis (African clawed frog)

C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000

C:Accession: B48088

R:Spevak, W.; Keiper, B.D.; Stralowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: B48088; MUID:93330289

A:Accession: B48088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SP>

A:Cross-references: GB:M88268; NID:9295542; PIDN:AAA02810.1; PID:929543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

| Query Match | 85.6% | Score 2597 | DB 2 | Length 518 |
|-----------------------|----------------|--|-----------|------------|
| Best Local Similarity | 91.6% | Pred. No. 1.6e-193 | | |
| Matches 488 | Conservative 7 | Mismatches 8 | Indels 30 | Gaps 1 |
| QY | 18 | SSEREDCNNGEPPRKITPEKNSLRQYNSCARLCLNQETVCLASTAMKTENCVAKTIAN | 77 | |
| DB | 13 | ASEREDCNNDPEPKITTEKNTLRQ-----TKIAN | 42 | |
| QY | 78 | GTSSMTVPKQKRLASYEKEELCVKYPEQWSESDQVEFVEHLISOMCHYQGHINTYK | 137 | |
| DB | 43 | GTSSMTVPKQKRLASYEKEELCVKYPEQWSESDQVEFVEHLISOMCHYQGHINTYK | 102 | |
| QY | 138 | PMQGRFTALPARGLDHTIAENILSYDAKSLCAELVCKEYRYSQGLMKLIERMV | 197 | |
| DB | 103 | PMQGRFTALPARGLDHTIAENILSYDAKSLCAELVCKEYRYSQGLMKLIERMV | 162 | |
| QY | 198 | RIDSILMRGLAERRGQGYLFKNKPPDGNAPNSFFRATYLPKTIIDTETIESNMRCGRSL | 257 | |
| DB | 163 | RIDSILMRGLAERRGQGYLFKNKPPDGNAPNSFFRATYLPKTIIDTETIESNMRCGRSL | 222 | |
| QY | 258 | QRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKINDKNTLECKRILLTGHGTVCLQYDE | 317 | |
| DB | 223 | QRHCRSETSKGYVCLQYDDOKIVSGLRDNTIKINDKNTLECKRILLTGHGTVCLQYDE | 282 | |
| QY | 318 | RYITIGSSDSTYRVMDVNTGEMLNTLIHHCFAVLHFRNGMAYVCSKDRSTAVYDMSAP | 377 | |
| DB | 283 | RYITIGSSDSTYRVMDVNTGEMLNTLIHHCFAVLHFRNGMAYVCSKDRSTAVYDMSAP | 342 | |
| QY | 378 | TDTTLRLVYVGRHAAVNVDFDOKIVSASGDRITIKVWNTSTCEFEVRLTNGHKRIGIACQ | 437 | |
| DB | 343 | TDTTLRLVYVGRHAAVNVDFDOKIVSASGDRITIKVWNTSTCEFEVRLTNGHKRIGIACQ | 402 | |

| Accession | Protein | Length |
|-----------|---|--------|
| QY438 | YRDLVYSSGSDNITRLMDIECGACLRVEGHEHLCIRCFPNKRIYSGAYGSKIKWML | 497 |
| Db403 | YRDLVYSSGSDNITRLMDIECGACLRVEGHEHLCIRCFEDPKRIYSGAYGSKIKWML | 462 |
| QY498 | VVALDPRAPAGTICLRILVHSGVAPFLQDEFOIYSSSHDITLLIWLDELNDP | 550 |
| Db463 | VVALDPRAPAGTICLRILVHSGVAPFLQDEFOIYSSSHDITLLIWLDELNDP | 515 |

RESULT

2

n/reboogall, k/c/v, najaunalcam, m.a.r., ballietl, b.g.; brown, s.; murphy, l.; jones, l.

submitted to the EMBL Data Library, January 2000

| | | | | |
|-----------------------|------------------|---------------------|------------|-------------|
| Query Match | 53.98; | Score 1635.5; | DB 2; | Length 701; |
| Best Local Similarity | 57.78; | Pred. No. 8.6e-119; | | |
| Matches 328; | Conservative 69; | Mismatches 116; | Indels 55; | Gaps 8 |

Db 527 SHDDTILWDFLDAP-----PSGLPST 549

RESULT 3

T50211

WD-repeat protein [Imported] - fission yeast (*Schizosaccharomyces pombe*)

C:Species: Schizosaccharomyces pombe

C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000

C:Accession: T50211

R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L., et al.

submitted to the EMBL Data Library, January 2000

A:Reference number: 225046

A:Accession: T50211

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-506 <MCD>

A:Cross-references: EMBL.AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05

A:Experimental source: strain 972h(-); cosmid c30

C:Genetics:

A:Gene: SPAC29E6.01; SPDB:SPAC30.05

A:Map position: 1

A:Introns: 43/1; 74/3

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

| | | | | |
|-----------------------|-------------------|--------------------|-------------|-------------|
| Query Match | 22.7%; | Score 690; | DB 2; | Length 506; |
| Best Local Similarity | 30.4%; | Pred. No. 1.1e-45; | | |
| Matches 163; | Conservative 104; | Mismatches 167; | Indels 102; | Gaps 15; |

```

Db      51  S1SNEGVAAYNHNVS1FTDTEVEF- ---EEVS1RFVSFDOJDLCKK1MSKMKRL 106
QY      183  TSDGM1MKKL1-----ERNVR1DS1MRG-----LMRRMG-----21
Db      107  LEDBG1KALYMKGFVNEV1LEEFEMRTHKFPQRPENFLKON1IGPYGM1LPQ 166
QY      214  QYLFKNKPPDGNAPNPFSEFALYKPI1D1D1ETTES1MRCGRHS1QRHRS-----264
Db      167  QYF1-----DSNGRPLM1MSY1-----KEAH1D1SMNRGR1LYSFVFN1PS1RPAODF 21
QY      265  -E1SKGY1C1Q1DDOK1YSGL1NDNT1K1MDKNT1ECKR111TG1SGV1C1Q1DYDER--V1I 321
Db      218  R11L1DS1Y1CYQ1YDDE1M1VSG1SKDR1V1SW1VNS1RFL1YK1LGH1SG1SV1C1D1E1C1RR1N1LY 27
QY      322  TGSSD1SV1RW1V1VNGEM1NTL1IHCEAV1LH1R1NNG1M1V1TOS1KRS1IA1V1--D1AS1TD 379
Db      278  SGS1SD1T1I1RW1M1QNR1RPLK1YF1GHT1DW1L1GV1V1SENY1I1SS1RHD1TAR1W1MD1TSP1AE 337
QY      380  I1LR1V1LGH1RA1V1VVD1DDK--Y1VAS1AGD1RT1K1WMT1S1TCE1V1R1LNG1H1K1G1AC1Q 437
Db      338  ACM--H1V1R1GH1LAS1V1NS1V1YS1K1TG1L1Y1AS1D1R1L1FT1W1IT1TG1HC1R111H1H1Q1G1AC1Q 396
QY      438  Y1R1R1V1VSG1SS1DNT1R1MD1EC1G1AC1V1L1EG1HE1V1C1R1PDK1R1V1SG1AY1DG1K1KY1WD1 497
Db      397  Y1NK1F1V1VSG1SS1D1T1R1IF1E1AS1G1K1L1M1L1Q1GH1ED1L1RT1R1F1NDE1K1V1SG1GD1V1R1M1N- 455
QY      498  V1AL1D1R1A1V1GL1C1R1L1V1EH-----SR1V1R1LQ1DFD1EQ1V1SS1HDD1T1L1MD1F 546
Db      456  -----E1NTE1Q1C1V1L1H1NS1R1V1SG1LQ1F1H1R1R1I1A1CTH1S1E1LL1VNF 497

```


| | | | | |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match | 17.5%; | Score 531.5; | DB 2; | Length 650; |
| Best Local Similarity | 25.0%; | Pred. No. 3.1e-33; | | |
| Matches 14; | Conservative 76; | Mismatches 180; | Indels 177; | Gaps 11 |

| | | | |
|----|-----|---|-----|
| QY | 117 | VEHLISOMKHQYGHGHNLSLTKPMLORDFTALTARGLDHIAENITLSYDAKSTCAELVC | 176 |
| Db | 99 | LOGILISQLCEFPOLSEFSRVEALNKIDFTISALPV-----ELAQKVLCEYLDLTVSTKAQVS | 154 |
| QY | 177 | KEMVYRTSDGMJMKLIEEMVTRTDSLMRGLAERBGWGQYLFNNKP-----P | 222 |
| Db | 155 | QGMRTIADSDAVWVRKRCBCHVNRK-----CTKCGMGLPILERKKLIRNTTROIACKGP | 208 |
| QY | 223 | DCNA-----PPNSFYRA | 234 |
| Db | 209 | QGRVTEIADSHDSODRSVNOHGKRPAAAEFEEDPTLKRCQMAAEASKAVTGPRTSMKA | 268 |
| QY | 235 | LVPKTIIDJETTESMMWRCGRHSLQRIHCSESFTSKYVLCIQYDQKIVSLRDNITKIDWK | 294 |
| Db | 269 | VY-----RDRWOVSYMMKNSRYLSV--KGHENGTTCIOLDNINLATISYDTITKIWNI | 321 |
| QY | 295 | NTECKRILITGHTGSVLCIQYDERVILITGSSDSYTRAVMVDNTEGMLNTLIHBCFAVLHR | 354 |
| Db | 322 | ETEECIRTLVGHTAGIRALQFPDCKSLISSLDHTIKVMMWHNGECLSTTAHTDSVISH | 381 |
| QY | 355 | PANGAMVTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNV--VDFDKYIVSASGDRTI | 412 |
| Db | 382 | FDGHLLIASSSSPKTYKIFEDFNSEKTYCLK---GHSMDVNSTHYDIKSPVSSADDTTI | 437 |
| QY | 413 | KVMNVTSTGEFVFTLNGHKRGJ-----ACLYQRDLRVLS----- | 445 |
| Db | 438 | KIMDLDTROVITYTTEGHVGHVQOVLILPEYEPDEEVLNGASQDNODAMSVSSGGSGS | 497 |
| QY | 446 | -----GSD | 449 |
| Db | 498 | MSHAQIERAGSPGSHSSHNLLPSSLPGDDEDEVRLHYGSAPAFADERSRLPPEYFMTGGID | 557 |
| QY | 450 | NTRIMDIECGACLRVLEGHEELVNCIRFDNKRIVGAYDGKIKWMDIYVALLPRAPAGT | 509 |
| Db | 558 | STMRLMDAATGACLTTLTLEGHLEGVWSLAGDITRIVISGANDGVKWE-----PRSG | 608 |
| QY | 510 | LCLRTLVHSGVRFLQFDEQIVSSHDDPTLLIMDF | 546 |
| Db | 609 | KCDATYTHCGPVTVCGLSDSLMSGSDGJITRLHSF | 645 |

```

RESULT      7
T22703
hypothetical protein F55B12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22703
R:Sims, M.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19602
A:Accession: T22703
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-579 <MIL>
A:Cross-references: EMBL:Z79757; PIDD:CA02129.1; GSPDB:GM00023; CESP:F55B12.3
A:Experimental source: clone F55B12
C:Genetics:
A:Gene: CESP:F55B12.3

```

| | | | | |
|-----------------------|------------------|------------------|------------|-------------|
| Query Match | 17.1%; | Score 520; | DB 2; | Length 579; |
| Best Local Similarity | 28.8%; | Pred. No. 2e-32; | | |
| Matches 150; | Conservative 78; | Mismatches 222; | Indels 70; | Gaps 17; |

```

0Y 72 KTKJLANGSSMT-----VKQRK--LSAYEKEKEK-----CVYFQBDMSDDVE 115
Db 27 ESSYSSNGSSSSYMDKXLSRSSRFLQHKLDLSAPSKNNLDNPREVHLLYALFKDLSAEBQMD 86
QY 116 FVEHLISQMCYQHGHINSYKPMLODPITLAPRGDHTAENILSYLDAKSCIAAEVY 175
Db 87 AFTRLLQSSNMNTNIRQLAIIIEPHQDRFLSCLPV---ELGMKLIHLNYGLDYDLKVAQV 142
QY 176 CKEWIRVSDCMLMKTL--IERVV-----RDSJLMRGLAERBGWGYLFFKKRPDGNAP 227
Db 143 SKNMKLISEIDKIMKSLGVEEFKHHDPDTRYGTAMQGTALAAQ-----VTRPDHQP 195
QY 228 PN-SYTRALYKRIIODT-----ETIESNMFCGRHSIORIHQSETSKGYVCL 273
Db 196 CDLVNHRFLKQKFGEDIFERADKSRYLRAKIEKNMNPIMGSAV--LRHEDHVIYCM 254
QY 274 QYDDOKIYSGLRDNTIKIKMDKNLTCEKRLLIHTGTSVCLUYDE--RYIYSSSDSTYRV 331
Db 255 QIHDDVLYTSGDDMTLKWCLDKGEMVMTLYVGHGTQGVWTSQISQCGRIYVSGSDTRYKV 314
QY 332 WDVNTGEMKLNLLIHCCEAVLHLRFENNGMMVTCSCDRSIAVMDMASPTDITLRYVLVGHRA 391
Db 315 WSTVDSGLLHTLQHTSTVRCMAMAGSLIVYGRSDTYTLRWMDVDSGRHLA---TLHGHH 371
QY 392 AVNVVDFEDDKIVASAGDRITKVVNTSTCEVPRTLNGHKRGIALCLOYDR--LVVSGSSD 449
Db 372 AVRCVQDFGTTIVASGGVDYFKIKNAHTRGCIIRTLTGNNNVYSLIRESESIYCSGSLD 431
QY 450 NTRLMWDI---ECGACLRVEEGHEELVRCIRFDNKRIVASGYDKRIKYMWDLYALDPRAP 506
Db 432 TSIWMPDFTREGEQCVALLQHTSLTSGMQLRGNIIIVSCNADSHVRVMDI-----H 483
QY 507 AGTLCRLTIVEHSGRVPRLO--FDEFQIYSSSHDPTIILMD 545
Db 484 EGT--CVHMLSGHSAITSLQWFGGRNMYATSSDDGTIVKLV 522

```

RESULT 8
S62507
hypothetical tlp-asp repeat-containing protein - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: T38502; S62507
R:Jones, L.; Murphy, L.; McNeill, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandream
submitted to the EMBL Data Library, October 1995
A:Reference number: Z21798
A:Accession: T38502
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <?02>
A:Cross-references: EMBL:Z65625; NID:g1044926; PIDN:CAA91423.1; PID:g1044927; GSPDB:G
C:Genetics:
A:Map position: 1
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:18-51/Domain: WD repeat homology <WD1>
F:104-137/Domain: WD repeat homology <WD2>

| | | | | | |
|----|-----------------------|--|--------------------|------------|-------------|
| | Query Match | 17.1% | Score 519.5 | DB 2 | length 267; |
| | Best local Similarity | 38.0% | Pred. No. 7.5e-33; | | |
| | Matches 104; | Conservative 58; | Mismatches 83; | Indels 29; | Gaps 6 |
| QY | 286 | DNTIKMDKNTLECRRLTGHTGSVLCQDYER--VIITGSSDSTVRWVDVTNGEMINTLT | 343 | | |
| | : | :::: | ::: | : | : |
| Db | 1 | DRVSVAWDNSNFELFKLYTGGSGSVLCLDFCRNRNLLSLGSSGSTIILIMWRNRRLPKAY | 60 | | |
| | : | :::: | ::: | : | : |


```

Oy      344 IHHCVAHLHAFNNGMAYTCCKSDSIAYW--DMSPTDITLRVLVGHRAAVNVVDEDDK 401
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      61 FGHDTNVLGVVYSENYIISSSRDHTRARWRDAPSPACM-HVLRGLHASVNSVSQSSK 119
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     402 --YIVSAGDRTIKWMNSTCEFTVLTNGHKRGJACLQYRRLVVGSGSDMTIRLMDIEC 459
       ||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     120 TGLIYTASDRTLRTWDITTTGHCIRIHAGRGJACAGYNKKFIYSSGSDLTIRIFEASS 179
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     460 GACLCRVLEGEHELVCIRCIFDNKRRIYSGAYDGRIKWMLVAALDPRAPAGTLCLRTVEH- 518
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     180 GKILRMLOGHEDLITVFAPFNDEKIYSGYGDGTVRIMN-----FVTGEQHQC 224
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     519 -----SGRVRFLODFEFOIYSSSHDDTILIMDF 546
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     225 VLHNRNSRVFGLQFDHRRIIACHHSSELVWNF 258
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT          9
T18521
beta transducin-like protein - Podospora anserina
C:Species: Podospora anserina
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18521
R:Saupe, S.; Turcq, B.; Begueret, J.
Gene 162, 135-139, 1995
A>Title: A gene responsible for vegetative incompatibility in the fungus Podospora anserina
A:Reference number: Z18944; MUID:96009891
A>Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1356 <SAU>
A:Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1
C:Genetics:
A:Gene: het-e1
A:introns: 761/3

Query Match           15.0%; Score 455.5; DB 2; Length 1356;
Best Local Similarity 26.0%; Pred. No.6.9e-27;
Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

Oy      28 EBPRIKIPKNSLRQTYNSCAR-LCLNQETVCLASTAMK-----TENCYAKTKLANGTSS 81
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     581 EMPDDVDSDPELDELIVKLGSGFLIRERTVYFHQSADFLIGTASDKASKNASQEAPE 640
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy      82 MYPK-----ORKLSATYEKEK--LCYKYREOMSESDOYEFVPHL--ISQMCHY 127
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     641 LVPEPTIEDVSYIIHFWSLVNWSOKLRRDIYCLINAGFLIDNVRPDPPLATVRYSCLY 700
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     128 OHGHINSYLKP-----MLORD---FIT-----ALPARGLDHI--AEN 159
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     701 WIDLHLDVIYSISSKKWWHLDDGDGIHREFLTIKYLYWLALSLIRLP-EGINALROLES 759
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     160 ILSYLDAKSICAEALVCKEMYRVTSDGMLMKLIER--MVRTDISLMRGLAERRGGQYL 216
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     760 LGHTIRGRLIA--IVRDGYRF--ALSVMITEKAQLPAVTSAL-----V 800
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     217 FKXKPPDGNAFPNSFFRALY----PKIIDIEFIESNMCGRSL----- 257
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     801 F-----APDMSMIKKIFKKEPGMISTISYVEAEEMNACTQTLEHGSSSVLSVASFAD 852
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     258 -ORIHRS-----ETSKG-----VYCLQY--DPDKIVSGLRONTIKIMDK 294
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     853 GQRVASGSDDKTIKIMDTASGTQTGLEHGGSVWSVAASPDERAVASGSDDKTIKIMDA 912
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     295 NTLECKRIITGHTGSVLCLOY--DERVLIITGSSDSVTVRWVNTGEMLNTLIHHCEAVLAH 352
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     913 ASGTCTQTLEGHGGRVQSVAFSPDGORVASGSDDHTIKIMDAASGTCTQTLEHGGSVLS 972
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy     353 LRHN-NGMWYC-SKRSIAVMDMASPTDITLRVLVGHRAAVNVVDF--DDXYIYASG 408
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     973 VAFSPDGORVASGSDDKTIKIMDTASG--TCTQTLEHGGSVWSVAFSPDGORVASGSD 1029

```


A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-703 <JAL>
 A:Cross-references: EMBL:AF064515; NID:93293382; PIDN:AAC39496.1; PID:93293383
 C:Genetics:
 A:Gene: pop2; sudi; SPAC4D7.03
 A:Map position: 1
 C:Function:
 A:Description: required to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;
 Best Local Similarity 23.3%; Pred. No. 1e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

```

14 KFMNSERED--CNGEPPPKI-----IPEKNSLRQTVNSCARLCL---NETVCL 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 KFMNSERED--CNGEPPPKI-----IPEKNSLRQTVNSCARLCL---NETVCL 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 ASYAMTENCVAKTKLANGTSSMIVPKOKKLSASYEK-EKELCVKFEQMSQDVEEVE 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 NSISSNSDNFPPSPRY--DTSNVSPGSKPISDEDLNLSIVOTFEEDLPFGIO--STAF 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 HLISQCHYQGH-INSYLPMLQDPTTALPARGLDHIAENILSYLDAKSLCAAEVCK 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 FQLLRSCNRSKMLNECEPLKDIENLPF---SIVQSLILNDIHSFLSGRLVSP 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 EWRV---TSDGMLKTLIERNVRDSTLRGLAERGGQYLFKNKPPDGNAAP---N 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 TMRRLDVTSS--YKMF-----SLFGQINEND--KYANENLRPPFLND 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 SEFVALYPRKIIDIETIESNMRCGRHSIQR----- 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313 QISDQFPEIFK-----RHFLNRKRWLFPSIPSHLSPIHVPNMTISLLH 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 -----HCRSETSK-GVYCLQDQKIVSGLENDTIKIMDKNT 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 KDRITTSSTGTTQIHMAITGVLEARLEGGKGVAAVKTIENTVLVSGSIDKIVRAVNIETK 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 LECRRLTGHTGTVLCQY-----DERITIGSSDSRYRVWDVNTGEML 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 ACSTHIFRHSIIRCLSELVPSRLRHGVETVEDDQPIYVSGSDHILRW----- 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 NTLIHCEAVLHFRNNGMMVTCSDRSIAVDMASPTDITLRRVYGHRAAVNVDFD 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
473 -----KLPRKTPDPYLPDNTNSIDRWE-KNPYFV---HLLIGHTSVKRTISGY 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 KTVASASGRTIKVWNTSCFVRLNGHKGIAQLQTRD--RLVYSSSSDNTIRLMDIE 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 DILVSGSYDSSIRIRVSTGDELYLRGHSRLIYSLVEPERNICISGSMKSIKRWDL 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 GCACLRVLEGHEELVRCIRFDPNKRIVSGAYDGKIVWDLVLAALDPRRAGTLCRTLVEH 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TGTCTVYLEGHDAFATLLANVQNRLISSASDSTRIMD-----LMTGPSN 628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 SGVFRLOQDEFQIVYSSHDITILMD 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
629 SGYISSFVSDEHKIT-SCNDGSKVLMD 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

S36113
 LIS-1 protein - human
 C:Species: Homo sapiens (man)
 C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 26-May-2000
 C:Accession: S36113
 R:Reiner, O.; Carozzo, R.; Shen, Y.; Wehnert, M.; Faustiniella, F.; Dobyns, W.B.; Caskey
 Nature 364, 717-721, 1993
 A:Title: Isolation of a Miller-Dieker lissencephaly gene containing a G protein beta-subunit
 A:Reference number: S36113; MUID:93361119
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-409 <RAS>
 C:Genetics:
 A:Gene: LIS-1
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:103-136/Domain: WD repeat homology <WD1>
 F:145-178/Domain: WD repeat homology <WD2>
 F:187-220/Domain: WD repeat homology <WD3>
 F:229-262/Domain: WD repeat homology <WD4>
 F:333-366/Domain: WD repeat homology <WD5>
 F:375-409/Domain: WD repeat homology <WD6>

Query Match 11.7%; Score 354; DB 2; Length 409;
 Best Local Similarity 29.2%; Pred. No. 9.3e-20;
 Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

```

280 IVSGLRDNTIKIMKNTLECKRLTGHTGTVLCQYDF--RVITGSSDSYRVWDVNTG 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 NVSASEDNITVWDETDGTFRTLKGDTSVQDISFHSGLKASCADMTIKLMDQGF 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 EMLNTLIHCEAVLHFR--NNGMMVTCSDRSIAVDMASPTDITLRRVYGHRAAVNV 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 ECIRTMGHGHDHNVSSVAIMPNGDHIVASRDKTIKMEVOTGYCV---KFTGHRVVRM 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
396 V--DPDDKYIVSAGDKRIKVNNTSCFVRLNGHKGIAQLQTRD----- 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 VRPNODGTLIASCNDQTVRVVAVATKECKALERHEHVEICISMAPESSYSISEATGS 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 -----LVVSGSSDNTIRLMDIEGACRLVLEGHEELVRCIRFDPN--KRIVSGAYD 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 ETKSGKRGPFLLSGSRKRTIKMDVSTGKMLTVLGHDMNVRVLFHSGKFTLSCADD 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
490 KRIWMDLVAALDPRRAGTLCRTLVEHSGRVFRLOQDEF--QIVSSSHDDITILMD 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
359 KTLRWMDY-----KNKRCMKTLNAHEHFTVSLDFHTAPVYVGSVDQTVKWE 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

S48052
 Platelet-activating factor acetylhydrolase 45k chain - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 26-May-2000
 C:Accession: S48052; S51606; S48075
 R:Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
 Nature 370, 216-218, 1994
 A:Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activat
 A:Reference number: S48052; MUID:94301405
 A:Accession: S48052
 A:Molecule type: mRNA
 A:Residues: 1-410 <HAT>
 A:Cross-references: EMBL:D30615
 A:Note: the title has been revised in S48075
 R:Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
 Nature 370, 391-391, 1994
 A:Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activat
 A:Reference number: S48075
 A:Contents: annotation
 A:Note: this is a revision to the title from reference S48052
 R:Hattori, M.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S51606
 A:Accession: S51606
 A:Molecule type: mRNA
 A:Residues: 1-25, 'A', 27-153, 'E', 155-410 <HAM>
 A:Cross-references: EMBL:D30615; NID:9516665; PIDN:BA06305.1; PID:d1006869; PID:9551
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:104-137/Domain: WD repeat homology <WD1>
 F:146-179/Domain: WD repeat homology <WD2>
 F:188-221/Domain: WD repeat homology <WD3>
 F:230-263/Domain: WD repeat homology <WD4>
 F:289-325/Domain: WD repeat homology <WD5>
 F:334-367/Domain: WD repeat homology <WD6>
 F:376-409/Domain: WD repeat homology <WD7>

| | | | | |
|-----------------------|--------|--------------------|--------|-----------------|
| Query Match | 11.7%; | Score 354; | DB 2; | Length 410; |
| Best Local Similarity | 29.2%; | Pred. No. 9.4e-20; | | |
| Matches | 87; | Conservative | 57; | Mismatches 110; |
| | | | Indels | 44; |
| | | | Gaps | 8; |

```

QY 280 IVSDBRNTKIMDKNLTLECKRILLTGTSVLQDYE--RVIITSSDSSTVAWVYNTG 337
      ||| ||| ||| : : : ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| |||
Db 123 MVSSEDATIKWMEYETGDERLLKHTTSDVODISPHSKLLASCADMTIKLMDPQGF 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 EMLNTLLIHCEAVLHLP--NGMAMVYCSKDRSIAWMDMASPTDITLRLVGHRAAVY 395
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 183 ECIPTMGHDHNVSSVAIMPNGDHIVASBDRKTIKMEVOTGYCV--KTFPGHREVMYR 239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 V--PEDDKYIVSAGSDATIKWYNTNCEEFRTLLGNHGRGLACLOYDR----- 441
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 VRPMQDGTLLASCNDQTVRWVYVATIECCAEELREHNVVECTSMAPESYSISENTGS 299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 -----LVYSSGSDNTRIIMDIECGACLRVLECHEELVHCIREDN--KRIVSGAYD 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 ETKSKSGRPFILGSRBDKTIKMDVSTGACIMTIIVGHDMVWGVLFHSGCKFLLSCADD 359
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 GKIVMDLVALLDRPAGTLCLETVIEHSGRVRLQDFEF--QIVSSSHDDIILMD 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 KTLRWMD-----KNRKQMTLANHEHVFISDFPHRTAPRVYVIGSDVQIVKWE 408
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

hypothetical protein YCR072c - yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Mar-1992 #sequence,revision 04-Dec-1992 #texLchange 26-May-2000
A:Accession: S19487; S26657
R:Bailesta, J.P.G.; Franco, L.; Hoenicke, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19486
A:Accession: S19487
A:Molecule type: DNA
A:Residues: 1-484, 'I', 486, 'TKL', 490, 492-515 <BAL1>
A:Cross-references: EMBL:X59720; MIPS:YCR072c
A:Note: this sequence has been revised in reference S26657
R:Bailesta, J.P.G.; Franco, L.; Hoenicke, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, October 1992
A:Reference number: S26587
A:Accession: S26657
A:Molecule type: DNA
A:Residues: 481-503 <BAL2>
A:Cross-references: EMBL:X59720; MIPS:YCR072c
C:Genetics:
A:Map position: 3R
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:182-215/Domain: WD repeat homology <WD2>
F:225-265/Domain: WD repeat homology <WD3>
F:274-306/Domain: WD repeat homology <WD4>
F:398-431/Domain: WD repeat homology <WD5>
F:440-473/Domain: WD repeat homology <WD6>

```

Query Match          11.2%;   Prec 341;   DB 2;   Length 515;
Best Local Similarity 30.8%;   Pred No. 1.3e+18;
Matches      89;   Conservative    47;   Mismatches    93;   Indels     60;   Gaps     9;

OY      256  STQRIHCRSESTKGVYCLQYDDOKIVSGLRDNTIKIMDKNLTCEKRILHTGTSVLCLOY 315
         ||| | | | : | : | | | | | : : | | | | | | : | : |
Db       236  SNEPRLHYVPKSG-----PRLASSCKDGITKIMDVYSVCQYTMSGHTNSVSCVM 286

OY      316  -DERVIITGGSDSYRVWVDVNF-GEMLNTLLIHCEAVLHLRFN-----356
         : : : | | | | | | : | : | : | : | : | : | : | : | : |
Db       287  GGQGLLYSGSHDRYVWDINSQGRICINLIKSHAHVNLSTLDYALRIGAFDHGTGKKP 346

OY      357  -----NG-----MMYTCKDSRIAAMDMASPDTILRRVLVIGHRA 392
         || | | | | | | | | : | : | : | : | : | : | : | : | : |
Db       347  STPEAKKALENYEICKRNKGENSEMAYTADDDTYTFIWNLPKSTPIAR--MGHQKL 404

```

OY 393 VNVVDE--DDKTVASGSRITKWVNSTCEBVTLLGNHRIAGLCLOYED--RLIVSSSS 448
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 405 VNHVASDPGRIVIASSDNSIKLMDGRDGKFISTFRGHIAASYGVAMASDPCRLVSCK 464
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 449 DNTIRLMDEGCACLRLEGGHEELRCRR--DNKRIVSAGVDGKRTKKV 495
| : | : | : | | | | | | | | | | | | | | | | | | | | | | |
Dd 465 DTTLAKWADVTRKLRSVLDLPGRHDEVYYDMSVDKRCSCSGDKDWKRWLM 513
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Search completed: May 8, 2002, 10:51:11
Job time: 123 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:48:33 ; Search time 25.03 Seconds
(without alignments)
1683.887 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034
Sequence: 1 MDPAEAVLQKALKEFNSS.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A.Geneseq.1101.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 3034 | 100.0 | 569 | 20 | AAV24054 |
| 2 | 3034 | 100.0 | 569 | 21 | AA12813 |
| 3 | 3034 | 100.0 | 569 | 21 | AAV96697 |
| 4 | 3034 | 100.0 | 569 | 21 | AAV83041 |
| 5 | 3034 | 100.0 | 569 | 21 | AAV83250 |
| 6 | 3034 | 100.0 | 569 | 21 | AAV44249 |
| 7 | 3027 | 99.8 | 569 | 22 | AA48298 |
| 8 | 2997 | 98.8 | 569 | 22 | AA12812 |
| 9 | 2997 | 98.8 | 569 | 21 | AA12812 |
| 10 | 2997 | 98.8 | 569 | 21 | AAV83254 |
| 11 | 2962 | 97.6 | 590 | 22 | AA100847 |

| | | | | | | |
|----|--------|------|-----|----|----------|---------------------|
| 12 | 2582.5 | 85.1 | 517 | 16 | AA85652 | WD-40 domain-contg |
| 13 | 2384.5 | 78.6 | 542 | 21 | AAV96696 | Human E3 ubiquitin |
| 14 | 2384.5 | 78.6 | 542 | 22 | AA40208 | Human polypeptide |
| 15 | 2384.5 | 78.6 | 550 | 22 | AA41994 | Human polypeptide |
| 16 | 935 | 30.8 | 448 | 22 | AA48289 | Beta-TRCP-N/SK2.C |
| 17 | 644.5 | 21.2 | 626 | 22 | AA52469 | Human gmyc-N-sel-1 |
| 18 | 644.5 | 21.2 | 626 | 22 | AA59201 | Protein encoded by |
| 19 | 640 | 21.1 | 540 | 20 | AAV22465 | Human hippocampal |
| 20 | 640 | 21.1 | 540 | 20 | AAV22468 | Human mammary sel- |
| 21 | 640 | 21.1 | 540 | 22 | AA59197 | Human hippocampal |
| 22 | 640 | 21.1 | 540 | 22 | AA59200 | Human mammary sel- |
| 23 | 640 | 21.1 | 545 | 22 | AA52464 | Human hippocampal |
| 24 | 640 | 21.1 | 545 | 22 | AA59196 | Human hippocampal |
| 25 | 640 | 21.1 | 545 | 22 | AA52463 | Human hippocampal |
| 26 | 640 | 21.1 | 533 | 22 | AA593475 | Human protein sequ |
| 27 | 640 | 21.1 | 553 | 22 | AA59195 | Human hippocampal |
| 28 | 640 | 21.1 | 559 | 22 | AAV22467 | Human mammary sel- |
| 29 | 640 | 21.1 | 559 | 22 | AA59199 | Human mammary sel- |
| 30 | 640 | 21.1 | 589 | 20 | AAV22466 | Human mammary sel- |
| 31 | 640 | 21.1 | 589 | 21 | AA501204 | Human GRPase assoc |
| 32 | 640 | 21.1 | 589 | 21 | AA59198 | Human mammary sel- |
| 33 | 640 | 21.1 | 592 | 20 | AAV22462 | Human hippocampal |
| 34 | 640 | 21.1 | 592 | 22 | AA59194 | Human hippocampal |
| 35 | 640 | 21.1 | 627 | 20 | AAV22461 | Human hippocampal |
| 36 | 640 | 21.1 | 627 | 22 | AA59193 | Human hippocampal |
| 37 | 640 | 21.1 | 666 | 20 | AAV22471 | Human C-term mychl |
| 38 | 640 | 21.1 | 666 | 22 | AA59203 | C-terminal mychl |
| 39 | 640 | 21.1 | 669 | 20 | AAV22470 | Human C-term V5 his |
| 40 | 640 | 21.1 | 669 | 22 | AA59202 | C-terminal V5 his |
| 41 | 545 | 18.0 | 640 | 21 | AAV83252 | F-box protein Met3 |
| 42 | 520 | 17.1 | 587 | 20 | AAV03204 | Amino acid sequenc |
| 43 | 399 | 13.2 | 779 | 16 | AA85854 | WD-40 domain-contg |
| 44 | 399 | 13.2 | 779 | 21 | AAV83249 | F-box protein Cdc4 |
| 45 | 396.5 | 13.1 | 703 | 21 | AAV83253 | F-box protein Pop2 |

ALIGNMENTS

| | | |
|----------|---|----------------------------|
| RESULT 1 | | |
| ID | AAV24054 | standard; Protein: 569 AA. |
| XX | | |
| AC | AAV24054; | |
| XX | | |
| DT | 30-SEP-1999 | (first entry) |
| XX | | |
| DE | A human beta-transducin repeat containing protein. | |
| XX | | |
| KW | Beta-transducin repeat containing protein; beta-Trcp; Skp1; | |
| KW | proteosome degradation pathway; Vpu protein; beta-actinin; | |
| KW | human immune deficiency virus-1; HIV-1; cellular protein; IkappaB; | |
| KW | ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's; | |
| KW | antiviral; antitumour; cell cycle regulation; protein degradation; | |
| KW | and anti-inflammatory; osteo-articular inflammation; acute inflammation; | |
| KW | tumour necrosis factor. | |
| OS | | |
| XX | | |
| XX | Homo sapiens. | |
| PH | | |
| FT | Key | Location/Qualifiers |
| FT | Region | 147..191 |
| FT | | /note= "F box sequence" |
| FT | Region | 259..292 |
| FT | | /note= "WD motif" |
| FT | Region | 304..332 |
| FT | | /note= "WD motif" |
| FT | Region | 343..372 |
| FT | | /note= "WD motif" |
| FT | Region | 387..415 |
| FT | | /note= "WD motif" |
| FT | Region | 427..455 |
| FT | | /note= "WD motif" |
| FT | Region | 467..492 |

```

FT      /note="WD motif"
FT      Region      516..544
FT      /note="WD motif"
XX      WC0938969-A1.
XX      PD
XX      05-AUG-1999.
XX      PF
XX      29-JAN-1999; 99MO-PR00196.
XX      PR
XX      09-DEC-1998; 98PR-0015545.
XX      30-JAN-1998; 98PR-0001100.
XX      PA
XX      (INRM ) INSEMT INST SANTE & RECH MEDICALE.
XX      PI
XX      Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;
XX      Kroll M, Margotin F;
XX      WPI; 1999-469329/39.
XX      DR
XX      N-PSDB; AAX86501.
XX      PT
XX      New human beta-transducin repeat containing protein and its
XX      fragments useful as, or to screen for, antiviral, antitumour,
XX      anti-inflammatory and anti-Alzheimer's agents
XX      PS
XX      Claim 1; Page 60-61; 71pp; French.
XX      CC
XX      The present sequence represents a human beta-transducin repeat containing
XX      protein (beta-Trcp). The protein directs proteins to the proteosome
XX      degradation pathways. The protein is able to interact with the Vpu
XX      protein of human immune deficiency virus-1 (HIV-1), cellular proteins
XX      Ikappab or beta-catenin (bc) and/or protein Skp1p. The protein controls
XX      ubiquitinylation of phosphorylated proteins and thus their targeting to
XX      proteosomes for degradation. Depending on whether the process is
XX      inhibited or promoted, the result may be delayed breakdown of CD4 (in
XX      cases of HIV-1 infection), increased activity of Ikb (and thus reduced
XX      activity of NFkappab) and increased degradation of mutant bc in tumour
XX      cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
XX      patients. The beta-Trcp protein, and its active peptide fragments, or its
XX      nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
XX      antitumour agents that disrupt cell cycle regulation or protein
XX      degradation in human tumour cells, and anti-inflammatory agents that
XX      disrupt activation by NFkappab. Fragments of the protein are also
XX      useful for treating osteo-articular inflammation or acute inflammation
XX      associated with release of tumour necrosis factor.
XX      SQ
XX      Sequence 569 AA:

```

```

Query Match      100.0%; Score 3034; DB 20; Length 569;
Best Local Similarity 100.0%; Pred. No. 3.2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      1 MDPAAVLOEKALKFMNSSEREDNNGEPRIITEKNSLQNTYNSCARLCLNGETVLA 60
XX      1 mdpaavloekalkfmnsserednngepriiteknslqntynsca rclclngetvla 60
XX      61 STAMKTENCVAKTKLANGTSSNIVPKORKLSASYEKEKELCVKFEQSESDOYEVEHL 120
XX      61 stamkTEncvaktKlangtssniVpKorklSAsyEKeKELcvKfEQseSDoyeVehL 120
XX      121 ISQMCHYGHGHNSTYKMLORDFTTALPANGLDHIAENILSYLDANSLSLCAELVCKEWY 180
XX      121 isqmchYghGhnstYkmlORdFTtAlPaNgldHiAenilSyldAnslSLcaelVckewY 180
XX      121 isqmchYghGhnstYkmlGRdftAlPaNgldHiAenilSyldAnslSLcaelVckewY 180
XX      181 RTSQGMILMKLIERVNTDSLMRGLARGMGOYLFPKNNKPPDGNAPNSRYALRYKIT 240
XX      181 rtsqgmilmkliervntdSLmrGLARgmGOylfPKnnKppdGNaPnsrYalrYkit 240
XX      181 rtsqgmilmkliervntdSLmrGLARgmGOylfPKnnKppdGNaPnsrYalrYkit 240
XX      241 QDIETIESNMKGRSLORICRSETSKGVYCLQYDDQKIVSGGLRDMTIRKIMDNTLECK 300
XX      241 qdiEtiesnmKgrSLORiCRsetSKgvYclQyDDqKiVsgGLrDMtIRkiMDntLeck 300
XX      241 qdiEtiesnmKgrSLORiCRsetSKgvYclQyDDqKiVsgGLrDMtIRkiMDntLeck 300

```

```

OY      301 RILTEGTSVLQYDERVITITGSSDSTVRWVDVNTGEMILTLIHCEAVLHLRFNNGM 360
DB      301 rilTEgtsvLqYdeRVITITgSSdStVRwVdVnTGEMiLTLIHceAVLHLrFNngm 360
OY      361 VTCSKDRSIANWDMASPDITLRLRVLVGHRANAVVDPDDKRYIASGDRITIKWNTSTC 420
DB      361 vtCSkdrSiAnwDMaSPdITlRLrVLvGHrANaVVdPDdKRYiASgDRITiKWntStc 420
OY      421 EEFRTLNCHKRGKICLOVRDLVYVSGSSDNTIRLMDIECGACLRVLEGHELVRCIRFDN 480
DB      421 eEfRtLnChKrgKiClOvRdLVyVsgSSdNtIRlMDIEcgAClRVlEGheLVrCiRfDn 480
OY      481 KRIYSGAYDGKIKYWDLVVALDPPRAGTCLRTLVHSGRVRFLQDFEQIYSSSHDT 540
DB      481 krIySgAYdgKiKyWdLVvAlDppRagTCLrTLvHsgRvRfLQdfEQiYssShdt 540
OY      541 ILIMDFLNDPAAQAEPPRSPRTYISR 569
DB      541 ilImDFlNDpAAQAEppRSPrtYIsr 569

```

RESULT 2

```

AAB12813
ID AAB12813 standard; protein; 569 AA.
XX
XX AAB12813:
XX
XX 27-NOV-2000 (first entry)
XX
XX

```

Human beta-transducin repeat containing protein (beta-Trcp) SEQ ID NO:3.

```

KW      Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
KW      beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
KW      gene therapy; colon cancer; beta-transducin repeat containing protein;
KW      beta-Trcp.
XX      OS
XX      Homo sapiens.
XX      PN
XX      JP2000166542-A.
XX      PD
XX      20-JUN-2000.
XX      PF
XX      02-DEC-1998; 98JP-0343437.
XX      PR
XX      02-DEC-1998; 98JP-0343437.
XX      PA
XX      (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX      DR
XX      WPI; 2000-485550/43.
XX      DR
XX      N-PSDB; AAA73132.
XX
XX      F-box protein of ubiquitin ligase SCF complex which promotes the
XX      ubiquitination of Ikappab or beta-catenin
XX      PS
XX      Claim 3; Page 10-12; 19pp; Japanese.
XX
XX      The present invention describes an F-box motif protein of ubiquitin
XX      ligase SCF complex which promotes the ubiquitination of Ikappab or
XX      beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
XX      complex (SCF complex) of F-box protein containing F-box motif and WD40
XX      repeat motif and has the amino acid sequence of 45 residues (AAB12811)
XX      or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
XX      ligase FWD1 protein) and (AAB12813, which is human beta-transducin
XX      repeat containing protein (beta-Trcp)). The F-box protein can be used for
XX      the gene therapy of colon cancer by being recombined to a virus vector.
XX
XX      Sequence 569 AA:

```

```

Query Match      100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 3.2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAAVLAQKALFEMNSEREDCNNGEPKRIPEKNSLRQYNSCARCLNOETVCLA 60
DB 1 mdpaaavlaqkalkfemnserecdngeppkkipexnslrqynscarlclngevcia 60
QY 61 STAKTEKCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQMSDDQVEFEVHL 120
DB 61 staktekcvaktklantssmivpkorklsasyekekelcvkyfeqmsddqvefvehl 120
QY 61 STAKTEKCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQMSDDQVEFEVHL 120
DB 61 staktekcvaktklantssmivpkorklsasyekekelcvkyfeqmsddqvefvehl 120
QY 121 ISOMCHYOHGINSYLRKMLORDFTALPARGLDHIENILSYDAKSLCAELVCKEWMY 180
DB 121 isomchygohghinsylrkmlordftalparglldhiienilsydaakslaaelvckewy 180
QY 181 RVTSDGMLMKKLIEMVRTDSLMRGLAERGMGOYLKKNPPDGNAPPNSFYRALYPKII 240
DB 181 rvtsgdgmllkklievmvrtdslmrgl aer gmgoylkknppdg nappnsfyralypkii 240
QY 241 QDIETIESNMRGRHSIORIRCRSETSKGVYCLQYDDOKIYSGLRDNTIKIMDKNTLECK 300
DB 241 qdietiesnwrgrhsiorircrsetskgyvclqyddokiysglr dntikimdkntleack 300
QY 301 RILTGHTGSVLCLQYDERVIITGSSDSTVYRWVDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 rilghtgsvlclqyderivitgssdstvyrwvdvntgemlntli hceavlh lrfnngmm 360
QY 361 VTCSKDRSIAVWDMASPDITLRVLVGHRAAVNVDPDKIYASGDRITIKWNTSTC 420
DB 361 vtcskdrsiavwdmaspditlr vlvghraavnvdpdkiyasgdr itikwnstc 420
QY 421 EHVRLTNGHKGRIACLOYRDLRVYSGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
DB 421 ehvrltngkhgriacloyrdlrv ysgssdntirlmdiegc lrvlegheelvrcirfdn 480
QY 481 KRIVSGAYDGKIKVWDVLAALDPRAAGTLCRLTVLHSGRVFLQDFEQIVSSSHDDT 540
DB 481 krivsgaydgkikvwdlvaa ldp rapagtlcrltlv lshgsvflrqdfefqvssshddt 540
QY 541 IILWDFLNDPAAQAEPPRSPTTYTISR 569
DB 541 iilwdf lndpaaqaepprspttytistr 569
RESULT 3
AA96697
ID AA96697 standard; Protein; 569 AA.
AC AA96697;
XX
XX 26-SEP-2000 (first entry)
XX
XX Human beta-TrCP.
XX
XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
XX nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
XX anti-inflammatory; immunosuppressive; cytostatic.
XX
XX Homo sapiens.
XX
XX WO200034447-A2.
XX
XX 15-JUN-2000
XX
XX 10-DEC-1999; 99WO-US293771.
XX
XX 10-DEC-1998; 98US-0210060.
XX
XX (SIGN-) SIGNAL PHARM. INC.
XX (YISS) YISSUM RES & DEV CO.
XX
XX Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzuba A;
XX Laron I, Yaron A;
XX
XX MPI; 2000-431294/37.

```

```

DR N-PDB; AAA51229.
XX
XX Polypeptide enhancing phosphorylated Ikbapab ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
XX Claim 21; Page 72-74; 77pp; English.
PS
XX
XX Human beta-TrCP, an F-box/WD protein family member, has been shown to
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of this process for use in treating diseases associated with
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX
XX Sequence 569 AA:
SQ
Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 3, 2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPAAVLAQKALFEMNSEREDCNNGEPKRIPEKNSLRQYNSCARCLNOETVCLA 60
DB 1 mdpaaavlaqkalkfemnserecdngeppkkipexnslrqynscarlclngevcia 60
QY 61 STAKTEKCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQMSDDQVEFEVHL 120
DB 61 staktekcvaktklantssmivpkorklsasyekekelcvkyfeqmsddqvefvehl 120
QY 121 ISOMCHYOHGINSYLRKMLORDFTALPARGLDHIENILSYDAKSLCAELVCKEWMY 180
DB 121 isomchygohghinsylrkmlordftalparglldhiienilsydaakslaaelvckewy 180
QY 181 RVTSDGMLMKKLIEMVRTDSLMRGLAERGMGOYLKKNPPDGNAPPNSFYRALYPKII 240
DB 181 rvtsgdgmllkklievmvrtdslmrgl aer gmgoylkknppdg nappnsfyralypkii 240
QY 241 QDIETIESNMRGRHSIORIRCRSETSKGVYCLQYDDOKIYSGLRDNTIKIMDKNTLECK 300
DB 241 qdietiesnwrgrhsiorircrsetskgyvclqyddokiysglr dntikimdkntleack 300
QY 301 RILTGHTGSVLCLQYDERVIITGSSDSTVYRWVDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 rilghtgsvlclqyderivitgssdstvyrwvdvntgemlntli hceavlh lrfnngmm 360
QY 361 VTCSKDRSIAVWDMASPDITLRVLVGHRAAVNVDPDKIYASGDRITIKWNTSTC 420
DB 361 vtcskdrsiavwdmaspditlr vlvghraavnvdpdkiyasgdr itikwnstc 420
QY 421 EHVRLTNGHKGRIACLOYRDLRVYSGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480
DB 421 ehvrltngkhgriacloyrdlrv ysgssdntirlmdiegc lrvlegheelvrcirfdn 480
QY 481 KRIVSGAYDGKIKVWDVLAALDPRAAGTLCRLTVLHSGRVFLQDFEQIVSSSHDDT 540
DB 481 krivsgaydgkikvwdlvaa ldp rapagtlcrltlv lshgsvflrqdfefqvssshddt 540
QY 541 IILWDFLNDPAAQAEPPRSPTTYTISR 569
DB 541 iilwdf lndpaaqaepprspttytistr 569
RESULT 4

```

AAV83041
 ID AAV83041 standard; Protein; 569 AA.
 AC AAV83041;
 XX
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FBP-1.
 XX
 XX F-box protein; FBP; diagnosis; treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 XX inflammatory disorder; human.
 OS Homo sapiens.
 XX
 PN WO200012679-A1.
 PD 09-MAR-2000.
 XX
 XX 27-AUG-1999; 99WO-US19560.
 PF
 XX 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 98US-0118568.
 PR 15-MAR-1999; 99US-0124449.
 XX
 XX (UYNY) UNITV NEW YORK STATE.
 PA
 XX
 PI Chlaur DS, Pagano M, Latres E;
 XX
 DR WPI: 2000-256635/22.
 DR N-PSDB; AA293350.
 XX
 PT Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases
 XX
 PS Disclosure; Figure 3a; 245pp; English.
 XX
 XX Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also major opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 XX
 XX Sequence 569 AA;
 SQ
 Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 3.2e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPEAVVQAEKALFEMNSEREDNNGEPRIKIIPEKNSLRORYNSCARCLNOETVCL 60
 DB 1 mdpeavvqekalkfemsereednngpeprkiipeknslrqynscarcldnqetvcla 60
 QY 61 STAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVYFEQWSESDOVEVEHL 120
 DB 61 stamktenvcvaktklangtssmivpkqklsasyekekelcvkyfegwssdqvetevhl 120
 QY 121 ISOMCHVOHGHINSYKPMQROFITLAPARGDHTAENLSTLDKSLCAALVCWEW 180
 DB 121 isgmchvohghinsylkpmqrofitlpargdhtaenlstldkslcaaelvcwekw 180
 QY 181 RVTSQGMWKKLLERWRTDSLMRGLAERGMQYLFKNKPPGNNAPPNSFYALPKIT 240

DB 181 rvtsqgmwkkllermwrtdslmrgl aer gwyalfknkppcgna pnsfyaly pkit 240
 QY 241 QDIETIESNWRGCRHSIORICHSETSKGVCLQYDQKIVSGLRDNTKIMDKNTLECK 300
 DB 241 qdietiesnwrgrhsiorichse ts k g v c l q y d q k i v s g l r d n t k i m d k n t l e c k 300
 QY 301 RILGHTGSVLCI QYDERVITITGSSDSYVWVDVNTGEMINTLIHCEAVLHIRENNGM 360
 DB 301 rilghtgs v l c i q y d e r v i t i t g s s d s y v w v d v n t g e m i n t l i h c e a v l h i r e n n g m 360
 QY 361 VTCSDRSIWMVMASTPTDLRRVVGHRAAVNVDPDKYIVSASGRTIKVMTSTNC 420
 DB 361 v t c s d r s i w m v m a s t p t d l r r v v g h r a a v n v d p d k y i v s a s g r t i k v m t s t n c 420
 QY 421 EFWRTLNGHRGIA CLQYRDRLVYSSSDNTIRLMDIEGACLRVLEGHEELVRCIRPDN 480
 DB 421 e f w r t l n g h r g i a c l q y r d r l v y s s s d n t i r l m d i e g a c l r v l e g h e e l v r c i r p d n 480
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGVRFLQFDEFQIVSSHDT 540
 DB 481 k r i v s g a y d g k i k v m d l v v a l d p r a p a g t l c r t l v e s h g v r f l q f d e f q i v s s h d t 540
 QY 541 ILIWFELNDPAAQAEPPRSPSRITYYISR 569
 DB 541 i l i w f e l n d p a a q a e p p r s p s r i t y y i s r 569
 RESULT 5
 AAV83250
 ID AAV83250 standard; Protein; 569 AA.
 AC AAV83250;
 XX
 XX 16-AUG-2000 (first entry)
 DT
 XX F-box protein hBetratCp.
 DE
 XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200022110-A2.
 PD 20-APR-2000.
 XX
 XX 08-OCT-1999; 99WO-US23705.
 PF
 XX 09-OCT-1998; 98US-0103787.
 PR
 XX (HARD) HARVARD COLLEGE.
 PA
 XX Zhou P, Howley P;
 PI
 XX WPI: 2000-317970/27.
 DR N-PSDB; AA293710.
 DR
 XX Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound
 XX
 XX Claim 9; Page 171; 185pp; English.
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or

XX Human cell signaling proteins useful for, e.g. diagnosing cell
 PT proliferative and inflammatory disorders -
 PS Claim 1; Page 77-78; 90pp; English.

CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded
 CC by cDNA obtained from Incyte clone 3239149 of COLACOL library. It is
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is
 CC found to be homologous to beta-transducin repeats containing
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as
 CC hybridisation probe for detecting CSIGP related sequences or allelic
 CC variants. Recombinant CSIGP can be produced in host cells by transforming
 CC them with genetically engineered vectors. Agonists or antagonists can be
 CC used in the treatment of cell proliferative and inflammatory disorders
 CC associated with decreased or increased CSIGP expression. CSIGP is used in
 CC the diagnosis, prevention and treatment of cell proliferative disorders
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.

XX Sequence 569 AA:

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 3.2e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNNSSREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 mdpaeavlgkalkfnnssredcngpprkiipeknslrqtynsarcclnqetvcla 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORKLSASYEKEKELCVYFPQWSESDQVEVEHL 120
 DB 61 stamkencvaktklngssmtvvpkorklsasyekelcvyfpqwsesdqvetvehl 120
 QY 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHAENTLSYLAKSLCAAEVCKEWY 180
 DB 121 isgmchyghhinsylkplmqdrfitalparglidhaentlsyldakslcaaeivckewy 180
 QY 181 RYTSDDGMLMKKLLIERVRRDLSLRGLAERMGQOYLFKNRPPGNNAPNSFYALPKII 240
 DB 181 rvtsgdgmllwkkllermvrrdlsrlrglaerrwgqyliknppdgnapnsfyalyplkii 240
 QY 241 ODITETESNMRCGRHSLOIRHCRSETSKGYCYQYDOKIVSGLRDNTIKIWDKNLECK 300
 DB 241 qditetlesnmrcgrhsloirhcrsetskgycclyyddqkivsglrdntlikwknleack 300
 QY 301 RIITGHTGSVLCQYDERVITIGSSDSTYRVWDVNTGEMLNTLIHHC EAVLHLRFNNGMM 360
 DB 301 riltghtgsvlclqyderivitigssdstyrvwdvntgemlntlihhceavlhrlfrnngmm 360
 QY 361 VYCSKRSIAVWMDASPTDITLRVLVGHRAAVNVDFDKYIVASGSDRTIVWNTSTC 420
 DB 361 vycskrsiavwmdasptdiltlrvlvghraavnvdfdkyivsaagdtlkwntstc 420
 QY 421 EFWRTNGHRCAGTACQYRDRLVSSSDNTIRLMDIEGACRLVLEGHELVRCIAEFDN 480
 DB 421 efwrtlnghrctagclqyrdrlvsssdntirlmdiegacrlvleghelevrciaefdn 480
 QY 481 KRIYSGAYDGKIKVMDLVLAALDPRAPAGTCLRTLYEHSRVRRLQDFEQIYSSSHDT 540
 DB 481 krtvsgaydgkikvmdlvlaaldrapagclclrtlyehsgrvrirlqdfeqivssshdt 540
 QY 541 ILIWDPLNDPAAQAEPPRSPRTYTYISR 569
 DB 541 illwplndpaaqaepprsprtytyistr 569

RESULT 7
 AAB48298
 ID AAB48298 standard; protein; 569 AA.
 XX
 AC AAB48298;

XX 02-APR-2001 (first entry)
 DT
 XX Human ZF11 protein.

XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 KW Bad; Bcl-2; tumour; cytostatic.

XX Homo sapiens.

XX WO200075184-A1.

XX 14-DEC-2000.

XX 05-JUN-2000; 2000MO-US15449.

XX 04-JUN-1999; 99US-0137494.

XX (UYYA) UNIV YALE.

XX Zhang H, Tsvetkov LM, Kondo T;

XX WPI: 2001-061703/07.

XX N-PSDB; AAC84610.

XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 PT involves altering levels of proteins such as S-phase kinase associated
 XX proteins 1, 2 and cullin/CDC53 proteins -

XX Claim 3; Page 130-132; 162pp; English.

XX The invention relates to methods of altering the polypeptide levels in a
 CC cell, using proteins selected from S-phase kinase associated proteins 1
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 CC cullin/CDC53 family of proteins). The method is useful for altering the
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
 CC that modulate interactions between SKP and target proteins are useful for
 CC treating tumours.

XX Sequence 569 AA:

Query Match 99.8%; Score 3027; DB 22; Length 569;
 Best Local Similarity 99.8%; Pred. No. 1.6e-286;
 Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNNSSREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 mdpaeavlgkalkfnnssredcngpprkiipeknslrqtynsarcclnqetvcla 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORKLSASYEKEKELCVYFPQWSESDQVEVEHL 120
 DB 61 stamkencvaktklngssmtvvpkorklsasyekelcvyfpqwsesdqvetvehl 120
 QY 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHAENTLSYLAKSLCAAEVCKEWY 180
 DB 121 isgmchyghhinsylkplmqdrfitalparglidhaentlsyldakslcaaeivckewy 180
 QY 181 RYTSDDGMLMKKLLIERVRRDLSLRGLAERMGQOYLFKNRPPGNNAPNSFYALPKII 240
 DB 181 rvtsgdgmllwkkllermvrrdlsrlrglaerrwgqyliknppdgnapnsfyalyplkii 240
 QY 241 ODITETESNMRCGRHSLOIRHCRSETSKGYCYQYDOKIVSGLRDNTIKIWDKNLECK 300
 DB 241 qditetlesnmrcgrhsloirhcrsetskgycclyyddqkivsglrdntlikwknleack 300
 QY 301 RIITGHTGSVLCQYDERVITIGSSDSTYRVWDVNTGEMLNTLIHHC EAVLHLRFNNGMM 360
 DB 301 riltghtgsvlclqyderivitigssdstyrvwdvntgemlntlihhceavlhrlfrnngmm 360

QY 361 VTCSKDRSIAVWDASPDITLRLVGVHRAVNVDPDDKIVSASGDRITKWNSTC 420
 |||||
 Db 361 vtcskdrslawdmaspditlrlrvlgvhraavnvdddkivysasgdrlikwnstlc 420
 QY 421 EFVRLNGHKGRIACLOLRDLVSGSSDNTIRLMDIEGACLRVLEGEELVRCIRFDN 480
 |||||
 Db 421 efvrlngkhrgiaclgyrdrlvsgssdntlrlwdiecgacrlvlegheelvrcirfdn 480
 QY 481 KRIVSGAYDGRIKIVWDIVAALDPRAPAGTLCRLTVHSGRVFRLOPDEFOIVSSSHDT 540
 |||||
 Db 481 krivsgaydgrikivwdlvaaldrpapagtlcrltlvehsgvfrlqfdefqvssshdt 540
 QY 541 ILIMFLNDPAAQAEPPSPSRTTYISR 569
 |||||
 Db 541 ilimflndpaaqaepprsprtylisy 569
 RESULT 8
 AAM00960
 ID AAM00960 standard; Protein: 608 AA.
 XX
 AC AAM00960;
 XX
 DT 01-OCT-2001 (first entry)
 DE Human bone marrow protein, SEQ ID NO: 436.
 XX
 DE Human: bone marrow; antiinflammatory; cytostatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153453-A2.
 XX
 PD 26-JUL-2001.
 PD
 PF 23-DEC-2000; 2000WO-US34960.
 XX
 PR 21-JAN-2000; 2000US-0486725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX
 PA (HYSE-) HYSBO INC.
 XX
 PI Ford JE, Boyle BJ, Tang YF, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX
 DR MPI; 2001-488707/53.
 DR N-PSDB; AAH90079.
 XX
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 XX
 PS Claim 10; Page 523-524; 648bp; English.
 XX
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence

CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 SQ Sequence 608 AA;
 QY
 Query Match 99.8%; Score 3027; DB 22; Length 608;
 Best Local Similarity 99.6%; Pred. No. 1,7e-286;
 Matches 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 40 mdpeavlyqekalkfmsnsredcngpeppkllpeknslrqvnscaiclnqetvcla 99
 QY 61 STAMKTEVCVAKTKLANGTSMIVPKORLSASYEKEELCVKFEQWSDQVEFEHL 120
 |||||
 Db 100 stamktenvcavaktklangtsmivpkqrksaeyekelcvkfeqwsdqvefehl 159
 QY 121 ISOMCHYOHGINSYLYKPMLODFITALPARGDLHIAENILSYLDAKSLCAELVCKEWY 180
 |||||
 Db 160 isgmchyghhinsylkpmldqdfitalpargldhiaenilslaydaksicaelvcckewy 219
 QY 181 RYTSQGLMKRLIERAVRTDSLMRGLAERKGGQYLRKKNRPPGKNAPNSFYRLYPKII 240
 |||||
 Db 220 rlytsqglmkrlieravrtdslmrgl aerkggylrkknppdgknapsfyalypkii 279
 QY 241 ODIEFIESNMWCGRHSIORICRSETSKGVYCLQYDOOKIVSGJRDNTIKIMDKNTLECK 300
 |||||
 Db 280 qdiefiesnmwcginsigrincrsetskgyyclqyddkivsgjrdntlkdntleck 339
 QY 301 RILTGHTSVLCLQYDERVITIGSSDSTVRWVDVNTGEMTLIHCEAVLHFRNNGM 360
 |||||
 Db 340 rilgtgsvlclqyderivitigssdstvrwvdvntgemtlilhceavlhfrnngm 399
 QY 361 VTCSKDRSIAVWDASPDITLRLVGVHRAVNVDPDDKIVSASGDRITKWNSTC 420
 |||||
 Db 400 vtcskdrslawdmaspditlrlrvlgvhraavnvdddkivysasgdrlikwnstlc 459
 QY 421 EFVRLNGHKGRIACLOLRDLVSGSSDNTIRLMDIEGACLRVLEGEELVRCIRFDN 480
 |||||
 Db 460 efvrlngkhrgiaclgyrdrlvsgssdntlrlwdiecgacrlvlegheelvrcirfdn 519
 QY 481 KRIVSGAYDGRIKIVWDIVAALDPRAPAGTLCRLTVHSGRVFRLOPDEFOIVSSSHDT 540
 |||||
 Db 520 krivsgaydgrikivwdlvaaldrpapagtlcrltlvehsgvfrlqfdefqvssshdt 579
 QY 541 ILIMFLNDPAAQAEPPSPSRTTYISR 569
 |||||
 Db 580 ilimflndpaaqaepprsprtylisy 608
 RESULT 9
 AAB12812
 ID AAB12812 standard; Protein: 569 AA.
 XX
 AC AAB12812;
 XX
 DT 27-NOV-2000 (first entry)
 DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
 XX
 DE Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikapab;
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;
 KW beta-TTCP.
 XX
 OS Mus musculus.
 XX
 PN JP200016542-A.
 XX

PD 20-JUN-2000.
 XX
 XX 02-DEC-1998; 98UP-0343437.
 PF
 XX 02-DEC-1998; 98UP-0343437.
 PR
 XX
 XX
 XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 PA
 XX WPI: 2000-485550/43.
 XX N-PSDB: AAA73131.
 DR
 XX
 XX
 PT F-box protein of ubiquitin ligase SCF complex which promotes the
 PT ubiquitination of IkappaB or beta-catenin
 CC
 PS Claim 2; Page 9-10; 19pp; Japanese.
 XX
 CC The present invention describes an F-box motif protein of ubiquitin
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a
 CC complex (SCF complex) of F-box protein containing F-box motif and Wd40
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
 CC ligase Fm1 protein) and (AAB12813, which is human beta-transducin
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for
 CC the gene therapy of colon cancer by being recombinant to a virus vector.
 CC
 XX
 SQ Sequence 569 AA:

Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 1.3e-283;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAAVIOEKLKFMNSSEEDCNNGEPPKIIPEKNSLMQTYNSCARCLDNOETVCLA 60
 DB 1 mdpaavioekalkfmnsseecnngeppkiipeknsliqlynscaclngetvclt 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORSLASYEKEKELCVKFEQMSSEDOYEFVHL 120
 DB 61 stamkencvakaklangssmtvpxkrlsasyekelcvkyfeqmsedoyefvehl 120
 QY 121 ISOMCHYOHGHNTSLKMLQORDFTTALPANGLDHIAENITSLYDANSCLCAELVCKEWMY 180
 DB 121 isomchynghnslkmlqordfttalparqldhlaenilslydaksclcaaelvckewy 180
 QY 181 RYTSQGMILMKLIERMVRTDSLWRGLARRRGMGOYLEFKNKPPDGNAPNSRYALYPKII 240
 DB 181 rvtsgmilmkliermvrtdslwrglaerrgqylfknkppdenapnsryalyphkii 240
 QY 241 ODIERIESNMKRCGRHSIORHCRSETSKGYCLOYDDOKIVSGLRDNTIKIWDNLTLECK 300
 DB 241 odieriesnmrcgrhsiorhcrsetskgycloyddokivsglrdntikikwdnltleck 300
 QY 301 RILNGHTGSVCLQYDERVITIGSSDSSTVYRWVDNTGEMINTLIHCEAVLHLRFNNGMM 360
 DB 301 rilnghtgsvclqyderivitigssdstvyrwvdnagemintlihceavlhrlfnngmm 360
 QY 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVVDKDYIVSASGDRITIKWNTSTC 420
 DB 361 vtcskdrsiavwdmasptditlrrvlvghraavnvvdtkdyivsagdrtikwnstc 420
 QY 421 EFVFTLNHKGKGIACLOYRDRLVYSSSDNTIRLMDIECGACLRVEGHEELVRCIRFDN 480
 DB 421 efvftlnhkgkiacloyrdrlvysssdntirlmdiecgacrlrvegheelvrcirfdn 480
 QY 481 KRIISGAVDGIKIVMDIYAAALDPRAPAGTLCRLTVHSGVFPLOJDEFOIYSSSHDDP 540
 DB 481 kriisgavdgikivmdiyaaaldprapagtlcrltlvhsghvfplojdefoiyssshddp 540
 QY 541 ILIWDPLNDPAAQAEPPRSPSRITYYSR 569
 DB 541 iliwdplndpaaqaepprspsrityysr 569

RESULT 10
 ID AAY83254
 AAAY83254 standard; Protein; 569 AA.

AC AAY83254;

DT 16-AUG-2000 (first entry)

DE F-box protein FWDlp.

XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW mouse; ss.

OS Mus musculus.

PN W0200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

XX (HARD) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI: 2000-317970/27.

DR N-PSDB: AA293714.

PT Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound

PS Claim 9; Page 184-185; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or
 CC increasing the level of a target polypeptide and for creating and
 CC expressing a destabilized polypeptide which is subjected to SCF
 CC mediated proteolysis. Degrading any desired protein in a cell is
 CC useful for preventing or treating diseases caused by the presence of
 CC abnormal amount of the specific polypeptides, for drug discovery and
 CC for gene therapy. Diseases treated include cancer, by degradation of
 CC oncoproteins, Huntington's disease, other proliferative disorders and
 CC microbial infections. The method provides a quick and easy
 CC alternative to gene knockout technology. The target polypeptide can
 CC be degraded at all stages, or a specific stage, of development in the
 CC mature animal.

SQ Sequence 569 AA:

Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 1.3e-283;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAAVIOEKLKFMNSSEEDCNNGEPPKIIPEKNSLMQTYNSCARCLDNOETVCLA 60
 DB 1 mdpaavioekalkfmnsseecnngeppkiipeknsliqlynscaclngetvclt 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORSLASYEKEKELCVKFEQMSSEDOYEFVHL 120
 DB 61 stamkencvakaklangssmtvpxkrlsasyekelcvkyfeqmsedoyefvehl 120

QY 121 ISOMCHYGHINSYKPMLOQDFITALPARGLDHIENIISYDAKSICAAELVCKEY 180
 XX
 PT 121 isgmchyghinsy kpm lqdfital par gldhi eni isy daksic aeelvckey 180
 Db 181 RYTSOSGLMKKILIERVNRDLSMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKII 240
 QY 181 RYTSOSGLMKKILIERVNRDLSMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKII 240
 Db 181 r y t s o s g l m k k i l i e r v n r d l s m r g l a e r g m g o y l f r k k p p d g n a p p s f y r a l y p k i i 240
 QY 241 ODIEETESNMRCGRHSIQRHCRSETSKGVCLQYDDOKIYVSGLRDNTIKIWDKNTLECK 300
 Db 241 o d i e e t e s n m r c g r h s i q r h c r s e t s k g v c l q y d d o k i y v s g l r d n t i k i w d k n t l e c k 300
 QY 241 qdletiesnmrcgrhs iqr hcrse tskgvclqyddok iysglrdntik iwdkntle ck 300
 Db 301 RIITGTSVLCQYDERYITIGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGM 360
 QY 301 RIITGTSVLCQYDERYITIGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGM 360
 Db 301 r i i t g t s v l c q y d e r y i t i g s s d s t y r v w d v n t g e m l n l i h c e a v l h r e n n n g m 360
 QY 361 VCSKSRSLAVMDMAPTDITLRVYVGHRAAVNVDPFDKTIYASGDRTIKWNMTSNC 420
 Db 361 v c s k s r s l a v m d m a p t d i t l r v y v g h r a a v n v d p f d k t i y a s g d r t i k w n m t s n c 420
 QY 421 EFVRTLNGHKGRIACLOYRDLRVVSGSSDNTIRLMDIEGACLRVLEGEELVRCIRFDN 480
 Db 421 e f v r t l n g h k g r i a c l o y r d l r v v s g s s d n t i r l m d i e g a c l r v l e g e e l v r c i r f d n 480
 QY 481 KRIVSGAYGKIKVMDLVALDPRAPAGTLCRLTVEHSGRFRLOFDEFOIYSSSHDT 540
 Db 481 k r i v s g a y g k i k v m d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d e f o i y s s s h d t 540
 QY 541 ILIMDFLNDPAQAEPSPRSRTTYISR 569
 Db 541 i l i m d f l n d p a q a e p s p r s r t t y i s r 569
 RESULT 11
 ID AAM00847 standard; Protein: 590 AA.
 XX
 AC AAM00847;
 DT 01-OCT-2001 (first entry)
 XX
 DE Human bone marrow protein, SEQ ID NO: 210.
 XX
 KW Human: bone marrow; antiinflammatory; cytosstatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153453-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 23-DEC-2000; 2000MO-US34960.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX
 PA (HYSE-) HYSO INC.
 XX
 PI Ford JE, Boyle BJ, Tang YF, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX
 DR WPI; 2001-488707/53.

DR N-PSDB: AAH89966.
 XX
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders
 XX
 PS Claim 10; Page 354-355; 648pp; English.
 XX
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 SQ Sequence 590 AA;
 Query Match 97.6%; Score 2962; DB 22; Length 590;
 Best Local Similarity 99.8%; Pred. No. 3.7e-280;
 Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 15 FNNSSEREDCNGNEPPRKTIPEKNSLRQTYNSCARLCLNCTVCLASTAMKENCVAKTK 74
 Db 15 f n n s s e r e d c n g n e p p r k t i p e k n s l r q t y n s c a r l c l n c t v c l a s t a m k e n c v a k t k 74
 QY 75 LANGTSSMIVPKQRKLSAYEKEKELCVKRFQWSESDQVEFVHLLISOMCHYGHINS 134
 Db 75 l a n g t s s m i v p k q r k l s a y e k e k e l c v k r f q w s e s d q v e f v h l l i s o m c h y g h i n s 134
 QY 96 LANGTSSMIVPKQRKLSAYEKEKELCVKRFQWSESDQVEFVHLLISOMCHYGHINS 155
 Db 96 l a n g t s s m i v p k q r k l s a y e k e k e l c v k r f q w s e s d q v e f v h l l i s o m c h y g h i n s 155
 QY 135 YLKPMLQORDFITALPARGLDHIENIISYDAKSICAAELVCKEYRVSQGMWKKII 194
 Db 135 y l k p m l q o r d f i t a l p a r g l d h i e n i i s y d a k s i c a a e l v c k e y r v s q g m w k k i i 194
 QY 156 YLKPMLQORDFITALPARGLDHIENIISYDAKSICAAELVCKEYRVSQGMWKKII 215
 Db 156 y l k p m l q o r d f i t a l p a r g l d h i e n i i s y d a k s i c a a e l v c k e y r v s q g m w k k i i 215
 QY 195 RMVFRDLSMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKIIODIEETESNMRCGR 254
 Db 195 r m v f r d l s m r g l a e r g m g o y l f r k k p p d g n a p p s f y r a l y p k i i o d i e e t e s n m r c g r 254
 QY 216 RMVFRDLSMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKIIODIEETESNMRCGR 275
 Db 216 r m v f r d l s m r g l a e r g m g o y l f r k k p p d g n a p p s f y r a l y p k i i o d i e e t e s n m r c g r 275
 QY 225 HSLQRIHCRSETSKGVCLQYDDOKIYVSGLRDNTIKIWDKNTLECKRIITGTSVLCLO 314
 Db 225 h s l q r i h c r s e t s k g v c l q y d d o k i y v s g l r d n t i k i w d k n t l e c k r i i t g t s v l c l o 314
 QY 276 HSLQRIHCRSETSKGVCLQYDDOKIYVSGLRDNTIKIWDKNTLECKRIITGTSVLCLO 335
 Db 276 h s l q r i h c r s e t s k g v c l q y d d o k i y v s g l r d n t i k i w d k n t l e c k r i i t g t s v l c l o 335
 QY 315 YDERYITIGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGMVCSDRSIAVWDM 374
 Db 315 y d e r y i t i g s s d s t y r v w d v n t g e m l n l i h c e a v l h r e n n n g m v c s d r s i a v w d m 374
 QY 336 YDERYITIGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGMVCSDRSIAVWDM 395
 Db 336 y d e r y i t i g s s d s t y r v w d v n t g e m l n l i h c e a v l h r e n n n g m v c s d r s i a v w d m 395
 QY 375 ASPFDITLRVYVGHRAAVNVDPFDKTIYASGDRTIKWNMTSNCFEVRLNKHKGRIA 434
 Db 375 a s p f d i t l r v y v g h r a a v n v d p f d k t i y a s g d r t i k w n m t s n c f e v r l n k h k g r i a 434
 QY 396 ASPFDITLRVYVGHRAAVNVDPFDKTIYASGDRTIKWNMTSNCFEVRLNKHKGRIA 455
 Db 396 a s p f d i t l r v y v g h r a a v n v d p f d k t i y a s g d r t i k w n m t s n c f e v r l n k h k g r i a 455
 QY 435 CLQYRDLRVVSGSSDNTIRLMDIEGACLRVLEGEELVRCIRFDNKRIVSGAYGKIKV 494
 Db 435 c l q y r d l r v v s g s s d n t i r l m d i e g a c l r v l e g e e l v r c i r f d n k r i v s g a y g k i k v 494
 QY 456 CLQYRDLRVVSGSSDNTIRLMDIEGACLRVLEGEELVRCIRFDNKRIVSGAYGKIKV 515
 Db 456 c l q y r d l r v v s g s s d n t i r l m d i e g a c l r v l e g e e l v r c i r f d n k r i v s g a y g k i k v 515
 QY 495 WDLVALDPRAPAGTLCRLTVEHSGRFRLOFDEFOIYSSSHDTILIMFNDPAAQA 554
 Db 495 w d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d e f o i y s s s h d t i l i m f n d p a a q a 554
 QY 516 WDLVALDPRAPAGTLCRLTVEHSGRFRLOFDEFOIYSSSHDTILIMFNDPAAQA 575
 Db 516 w d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d e f o i y s s s h d t i l i m f n d p a a q a 575
 QY 555 EPPRSRSPRTTYISR 569
 Db 555 e p p r s r s p r t t y i s r 569
 RESULT 12
 AAR85852
 ID AAR85852 standard; peptide: 517 AA.

```

XX AC AAR85852;
XX DT 13-SEP-1996 (first entry)
XX DE WD-40 domain-contg. beta-TRCP protein.
XX DE WD40 repeat region; beta-transducin; protein-protein interaction; drug;
XX DE intracellular signalling; protein kinase C; homology; motif; modulator;
XX DE receptors of activated protein kinase; enzyme activity; isozyme; human.
XX OS Synthetic.
XX PN MO9521252-A2.
XX PD 10-AUG-1995.
XX PF 31-JAN-1995; 95WO-US01210.
XX PR 01-FEB-1994; 94US-0190802.
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX PI Mochly-Rosen D, Ron D;
XX DR WPI; 1995-283772/37.
XX PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
XX PT activity of a protein, eg. protein kinase C, which interacts with a
XX PT protein contg. a WD-40 region.
XX PS Example 5; Page 80-82; 351pp; English.
XX CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also
XX CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
XX CC regions are involved in protein-protein interactions between proteins
XX CC involved in intracellular signalling. An example of such an interaction
XX CC is between protein kinase C and receptors of activated protein kinase
XX CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based
XX CC on homology with beta-transducin, whereas proteins AAR85882-92 were
XX CC isolated based on homology with the WD-40 consensus sequence (AAR85893).
XX CC The proteins were used to construct the peptides AAR84928-R85063 and
XX CC AAR85766-R85842. The peptides can be used to identify target proteins
XX CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
XX CC proteins involved in protein-protein interaction and to screen for drugs
XX CC that will affect protein-protein interaction involving WD-40 domains.
XX SO Sequence 517 AA:

Query Match 85.1%; Score 2582.5; DB 16; Length 517;
Best Local Similarity 91.4%; Pred. No. 3,4e-243;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

OY 18 SSERECNNGEPPRKIIPEKNSLRQYNSCARLCINQETVCLASIAAMTKMCAKTKLAN 77
DB :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13 aserecndprrkllleknllr-----tklan 42

OY 78 GTSSMIVKQRKLASYEKEKLCVYFEOWSESDVEFEHLISOMCHYOHGHTNSYLK 137
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 gtsmivkqrklasyekekclvkyfegsecdqefvehllismchghghnhtylk 102

OY 138 PMIQDFTTALPARGLDHAENILSYDAKSLCAAEVLCKEMRYRTSDGMLMKRLIERMV 197
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
103 pmiqdfittalparglidhaenilsydakslcsaelvckewyrvtsdgmllwkkliermv 162

OY 198 RFDLSMRGLAERRGKGOVLFRKMKPPDGNAAPNSFRALYKPTIODIETIESNMWCGRRSL 257
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 rfdlsmrgl aer rgkgo vlfrk mppdg naapns fral ykptiod iet iesnmw cgr rsl 222

OY 258 QRHCRSETSGKVYCLQYDDOKIYSGELRDNTIKIMDKNTLECKRKLTHGTSVCLQYDE 317
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 qrhcrsetsgkvyclqyddokiysgelrdntikimdkntleckrklthgtsvclqyde 282

```

```

OY 318 RVIIITGSSDSTFRVWDVNTGEMLTLLIHCEAVLHLRFNNGMAMVTCSDKRSIAVMDASP 377
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 rviitg-sdstfrvwdvntgemtlllhceaavlhlrfnngmmtcskdsiavwmdasa 341

OY 378 TDILRLRVLGHRAAVVWVPDDKYYIYSASGDRIRKWNSTGCFVITLNGHKGRIACLO 437
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
342 tdlrlrvlghraavvwdpdkyyiysasgdrirkwnstgcfvltlngthkrgriaclo 401

OY 438 YRDLRVVSGSSDNTIRLMDIECGACLRVLDGHEELVRCIFRDNKRRIYSGAYDGRIKRWDL 497
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 yrdlrvvsgssdntirlmdiecgacrlrvldgheelvrcifrdnkrrivsgaydgrikrvwdl 461

OY 498 VAALDPRAPACTLCLRLVHSGRVFRLQDFDEQIYSSSHDDPTLLMDFTLNDP 550
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 vaaldprapactlclrlvhsgrvfrlqdfdeqiyssshddptllmwfndp 514

RESULT 13
ID AAY96696 standard; Protein: 542 AA.
XX AC AAY96696;
XX DT 26-SEP-2000 (first entry)
XX DE Human E3 ubiquitin ligase.
XX DE E3 ubiquitin ligase; beta-TRCP; F-box; WD protein; I-kappa-B; inhibitor;
XX DE nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
XX DE anti-inflammatory; immunosuppressive; cytostatic.
XX OS Homo sapiens.
XX PN W0200034447-A2.
XX PD 15-JUN-2000.
XX PF 10-DEC-1999; 99WO-US29371.
XX PR 10-DEC-1998; 98US-0210060.
XX PA (SIGN-) SIGNAL PHARM INC.
XX PA (YISS ) YISSUM RES & DEV CO.
XX PI Manning AM, Mercurio F, Amlt S, Ben-eriah Y, Davis M, Hatzubai A;
XX PI Lavon I, Yaron A;
XX DR WPI; 2000-431294/37.
XX DR N-PSDB; AAA51228.
XX PT Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
XX PT treating disorder associated with NF-kappaB activation e.g. cancer,
XX PT compising amino acid sequence of human E3 ubiquitin ligase or its
XX PT variant
XX PS Claim 1; Page 70-72; 77pp; English.
XX CC This is human E3 ubiquitin ligase (E3), which is homologous to human
XX CC beta-TRCP, an F-box/WD protein family member. E3 enhances ubiquitination
XX CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor
XX CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the
XX CC ubiquitin pathway is useful for identifying modulators of this process
XX CC for use in treating diseases associated with activation of NF-kappa-B. In
XX CC vitro analysis suggests that deletion of the F-box results in a protein
XX CC that functions as a dominant negative molecule in vivo. Transient
XX CC over-expression of delta-beta-TRCP (a deletion mutant) inhibited the
XX CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,
XX CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be
XX CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TRCP
XX CC can be used to modulate NF-kappa-B to treat inflammatory diseases,
XX CC autoimmune diseases, cancer and viral infections.

```


Db 332 lmvtskdrslavwmasatdltrrvlyghraavnvddfdkkyivasgdrtlkwwsts 391
 QY 419 TCEFVRLNGHKKRIACIQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGEHELYRCIRF 478
 CC |||||||
 CC 392 tcefvrlngkhrgiaclqyrdrlvsvgsndtlrlwdiecgacrlvlegheelyrcirf 451
 CC |||||||
 QY 479 DNKRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHD 538
 Db 452 dnkrivsgaydgkikvwdlgaaldprapastlclrtlvehsgrvrlqdfefqivssshd 511
 QY 539 DTILIMDFLNDPAAQAEPSPSRRTYTYISR 569
 Db 512 dtlilwdflnvpsaqnetrsprtytyisr 542
 RESULT 15
 AAM41994
 ID AAM41994 standard; Protein: 550 AA.
 XX AAM41994;
 AC AAM41994;
 XX 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 6925.
 DE Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS WO200133312-A1.
 PN 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US34263.
 PF 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Huang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Dzmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AAI61150.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 6925; 10078P; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 550 AA:
 SQ
 QY 1 MDPRAEVLQKALKFNNSSREDCNNGEPPRKLIPEKNSLRQYNSCARLQNOEVCIA 60
 Db 9 mep-dsylvlecdkcltclms-----vp-----tslwgcanlv---esmcsl 44
 QY 61 S--TAMKTENCVAKTKLANGTSSMIVPKORKLSAYEKEKELCVKFEQMSSEDOVEYE 118
 Db 45 sclqsmprsvrcl---qngtssvsvsrkprsegnyqkcklcklyfdqwsedqyefve 101
 QY 119 HLISQMCYHGHINSLTKMLQRPFTALPAGLDIHANILSYLDAKSICAEIYCKE 178
 Db 102 hlismchghghlnsltkpmlqrdltalpeghldhlaenllsyldarslcaaelvcke 161
 QY 179 WYRTSDGMIMKRLIERMVRDTSIMRGLAERGMGYLFENKPPDGNAPPNSFYRALYPR 238
 Db 162 wgrysgmimkrlkiermvrtdpwwglseriywdqylfknrptdg--ppnsfyrsllypk 219
 QY 239 IIODIETIESNWRGSRHSRLRHCSETSKGVYCLQYDQKIVYSGLRDNTIKWKNITLE 298
 Db 220 iigdieltesnwrgrhnlgrdqrsenskyvclgyddeklisgirdnsiklwdktsle 279
 QY 299 CKRLTGTHTGTSVYCLQYDERVITITGSSDSTVYRWVDVNTGEMTLIHCEAVLHLPFNN 358
 Db 280 cikvlgtgtsvclqydervivgsdsctervwdvntgevtllhneavvlhlfsng 339
 QY 359 MMYVCSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDKRYIVSASGDRITKWNMTS 418
 Db 340 lmvtskdrslavwmasatdltrrvlyghraavnvddfdkkyivasgdrtlkwwsts 399
 QY 419 TCEFVRLNGHKKRIACIQYRDRLVYVSGSSDNTIRLMDIEGACIRVLEGEHELYRCIRF 478
 Db 400 tcefvrlngkhrgiaclqyrdrlvsvgsndtlrlwdiecgacrlvlegheelyrcirf 459
 QY 479 DNKRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHD 538
 Db 460 dnkrivsgaydgkikvwdlgaaldprapastlclrtlvehsgrvrlqdfefqivssshd 519
 QY 539 DTILIMDFLNDPAAQAEPSPSRRTYTYISR 569
 Db 520 dtlilwdflnvpsaqnetrsprtytyisr 550

Search completed: May 8, 2002, 10:50:24
 Job time: 111 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:48:53 ; Search time 14.24 Seconds
(without alignments)
899.184 Million cell updates/sec

Title: US-09-601-168A-2

Perfect score: 3034

Sequence: 1 MDPAEAVLQKALKFMNSSE.....PAAQAPPPSPRTYYSR 569

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|---------------------|-------------------|
| 1 | 2582.5 | 85.1 | 517 | 1 US-08-190-802A-30 | Sequence 30, Appl |
| 2 | 2582.5 | 85.1 | 517 | 4 US-08-477-346-30 | Sequence 30, Appl |
| 3 | 520 | 17.1 | 587 | 3 US-08-899-578-2 | Sequence 2, Appl |
| 4 | 399 | 13.2 | 779 | 1 US-08-190-802A-32 | Sequence 32, Appl |
| 5 | 399 | 13.2 | 779 | 4 US-08-477-346-32 | Sequence 32, Appl |
| 6 | 354 | 11.7 | 409 | 2 US-08-283-917-3 | Sequence 3, Appl |
| 7 | 354 | 11.7 | 409 | 2 US-08-961-716-3 | Sequence 3, Appl |
| 8 | 354 | 11.7 | 409 | 2 US-08-283-917-9 | Sequence 9, Appl |
| 9 | 354 | 11.7 | 409 | 2 US-08-961-716-9 | Sequence 9, Appl |
| 10 | 339.5 | 11.2 | 409 | 4 US-08-190-802A-51 | Sequence 51, Appl |
| 11 | 339.5 | 11.2 | 409 | 4 US-08-477-346-51 | Sequence 51, Appl |
| 12 | 321.5 | 10.6 | 514 | 1 US-08-190-802A-66 | Sequence 66, Appl |
| 13 | 321.5 | 10.6 | 514 | 4 US-08-477-346-66 | Sequence 66, Appl |
| 14 | 318 | 10.5 | 422 | 1 US-08-190-802A-52 | Sequence 52, Appl |
| 15 | 318 | 10.5 | 422 | 4 US-08-477-346-52 | Sequence 52, Appl |
| 16 | 313.5 | 10.3 | 209 | 3 US-08-899-578-6 | Sequence 6, Appl |
| 17 | 306 | 10.1 | 704 | 1 US-08-188-582-5 | Sequence 5, Appl |
| 18 | 306 | 10.1 | 704 | 1 US-08-646-715-5 | Sequence 5, Appl |
| 19 | 305 | 10.1 | 704 | 1 US-08-190-802A-62 | Sequence 62, Appl |
| 20 | 305 | 10.1 | 704 | 2 US-08-308-818-3 | Sequence 3, Appl |
| 21 | 305 | 10.1 | 704 | 4 US-08-477-346-62 | Sequence 62, Appl |
| 22 | 299.5 | 9.9 | 1194 | 4 US-09-092-508-2 | Sequence 2, Appl |
| 23 | 298.5 | 9.8 | 1205 | 4 US-09-092-508-16 | Sequence 16, Appl |
| 24 | 298.5 | 9.8 | 704 | 1 US-08-188-582-18 | Sequence 18, Appl |
| 25 | 298.5 | 9.8 | 704 | 1 US-08-646-715-18 | Sequence 18, Appl |
| 26 | 285.5 | 9.4 | 798 | 1 US-08-190-802A-64 | Sequence 64, Appl |
| 27 | 285.5 | 9.4 | 798 | 1 US-08-190-802A-68 | Sequence 68, Appl |

| | | | | | |
|----|-------|-----|-----|---------------------|-------------------|
| 28 | 285.5 | 9.4 | 798 | 2 US-08-308-818-2 | Sequence 2, Appl |
| 29 | 285.5 | 9.4 | 798 | 4 US-08-477-346-64 | Sequence 64, Appl |
| 30 | 285.5 | 9.4 | 798 | 4 US-08-477-346-68 | Sequence 68, Appl |
| 31 | 283.5 | 9.3 | 212 | 3 US-08-899-578-7 | Sequence 7, Appl |
| 32 | 283 | 9.3 | 713 | 1 US-08-190-802A-63 | Sequence 63, Appl |
| 33 | 283 | 9.3 | 713 | 4 US-08-477-346-63 | Sequence 63, Appl |
| 34 | 263 | 8.7 | 318 | 1 US-08-190-802A-33 | Sequence 33, Appl |
| 35 | 263 | 8.7 | 318 | 4 US-08-477-346-33 | Sequence 33, Appl |
| 36 | 260.5 | 8.6 | 375 | 4 US-09-063-743-1 | Sequence 31, Appl |
| 37 | 256 | 8.4 | 906 | 1 US-08-190-802A-31 | Sequence 31, Appl |
| 38 | 256 | 8.4 | 906 | 4 US-08-477-346-31 | Sequence 31, Appl |
| 39 | 246.5 | 8.1 | 343 | 4 US-09-063-743-5 | Sequence 5, Appl |
| 40 | 245.5 | 8.1 | 317 | 1 US-08-190-802A-27 | Sequence 27, Appl |
| 41 | 245.5 | 8.1 | 317 | 1 US-08-190-802A-41 | Sequence 41, Appl |
| 42 | 245.5 | 8.1 | 317 | 1 US-08-190-802A-47 | Sequence 47, Appl |
| 43 | 245.5 | 8.1 | 317 | 4 US-08-477-346-27 | Sequence 27, Appl |
| 44 | 245.5 | 8.1 | 317 | 4 US-08-477-346-41 | Sequence 41, Appl |
| 45 | 245.5 | 8.1 | 317 | 4 US-08-477-346-47 | Sequence 47, Appl |

ALIGNMENTS

RESULT 1
US-08-190-802A-30
Sequence 30, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 5.5e-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

RESULT 2
 US-08-477-346-30
 Sequence 30, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WO-40 - Derived Peptides and Uses
 TITLE OF INVENTION: thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morlison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763

| | | | | |
|---------------------------|--------|---------------------|------------|-------------|
| Query Match | 85.1%; | Score 2582.5; | DB 4; | Length 517; |
| Best Local Similarity | 91.4%; | Pred. No. 5.5e-266; | | |
| Matches 487; Conservative | 7; | Mismatches 8; | Indels 31; | Gaps 2 |

| | | | |
|----|-----|--|-----|
| QY | 18 | SSEEDONNGNEPPRKITPEKNSLRBTYNSCARCLNGEYCLASTAMKEKNCVATKTAN | 77 |
| Db | 13 | ASEREDONDEPPRKITTEKNTLRO-----TKLAN | 42 |
| QY | 78 | GTSMMIVPORKLSASYEKEKELCVKYEFEWSESQVEFVEHLISOMCHYQGHINSYLK | 137 |
| Db | 43 | GTSMMIVPORKLSANEKEKELCVKYEFEWSECDQVEFVEHLISMCHYQGHINTYLK | 102 |
| QY | 138 | PMLOPFTALPARGIDHIAHNLISYDASLSLAELVECKEMVRYTSDOMLKKLIEKV | 197 |
| Db | 103 | PMLOPFTALPARGIDHIAENILSYDASLSLAELVECKEMVRYTSDOMLKKLIEKV | 162 |
| QY | 198 | RTDSLWRGLAERBGQVLFKNKPPDGNAPNSFYALYPKIIODIETIESMWRGRSL | 257 |
| Db | 163 | RTDSLWRGLAERBGQVLFKNKPPDGTPPNSFYALYPKIIODIETIESMWRGRSL | 222 |
| QY | 258 | QRHCSEFSKVCYLOQYDDQKITVSGLRDNTIKIMDKNLTBECKRLIIGTGSVYLCLOYDE | 317 |
| Db | 223 | QRHCSEFSKVCYLOQYDDQKITVSGLRDNTIKIMDKNLTBECKRYLMLGTGSVYLCLOYDE | 282 |
| QY | 318 | RVLTIGSSDSITRVWVDVNTGMLNTLIIHCEAVLHLRFNNGMAYVCSKRSIAVWDAASP | 377 |
| Db | 283 | RVLTIGSSDSITRVWVDVNTGMLNTLIIHCEAVLHLRFNNGMAYVCSKRSIAVWDAASA | 344 |
| QY | 378 | TDTTLTRVLYGHRAAVNVYDDDKYIYASAGDRTIKVMWSTICEFVRYTLNHHKRGIAQLO | 437 |
| Db | 342 | TDTTLTRVLYGHRAAVNVYDDDKYIYASAGDRTIKVMWSTICEFVRYTLNHHKRGIAQLO | 401 |
| QY | 438 | YRDRLTVSSGSDNTIRLMDIECGALRYLGEHBEELVRCIRFPNKRITVSGAYDGRKIRYWDL | 497 |
| Db | 402 | YRDRLTVSSGSDNTIRLMDIECGALRYLGEHBEELVRCIRFPNKRITVSGAYDGRKIRYWDL | 461 |
| QY | 498 | VALADPRAPAGTLCRTLYVHSGVFPLOQDEQOYVSSSHDPTILINPFLNDP | 550 |
| Db | 462 | VALADPRAPAGTLCRTLYVHSGVFPLOQDEQOYVSSSHDPTILINPFLNDP | 514 |

RESULT 3
 US-08-899-578-2
 Sequence 2, Application US/08899578
 Patent No. 6087153
 GENERAL INFORMATION:
 APPLICANT: Greinwald, Iva
 APPLICANT: Hubbard, E. Jane
 TITLE OF INVENTION: 'SEL-10 AND USES THEREOF'
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,578
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53200/JFW/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-578-2

```

```

Query Match      17.1%; Score 520; DB 3; Length 587;
Best Local Similarity 28.8%; Pred. No. 2,3e-46;
Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

```

```

QY 72 KTKLANGSTSMI-----VPRKOR--LSASYEKEL-----CVKYFEQMSDQVE 115
DB 35 ESSYSSNGSSSYNADKSSRPLOKHLDLSASPSRNNNDLPVREHLIALFKDLSAEDM 94
QY 116 FVEHLISQMCYHGHINSYLPMLQDPTLAPARGDHAENILSYDAKSLCAAEV 175
DB 95 AFTRLDQSMNTNIRQLRAIEPHRFDFLCLPV---ELGKMLHLNLGYDILLKVAQV 150
QY 176 CKERYRTSDGMLMKL-IERM-----RTDSLMRGLAERRGQGYLFKNRPPDGNAP 227
DB 151 SKNKKLISEIDKIMKSLGVEEFKHHPDPTDRVTGAWOGTIAAG-----VILPHIOP 203
QY 228 PN-SFYALPKIIDI-----ETIESNRCGRHSIORHCRSETSKGYCL 273
DB 204 CDLAVHFKLKQKRGDIFERRADKSRRLADKIEKNANPIMGSAV-LGHEHDVITCM 262
QY 274 QYDQKIVSGLRDNTIKIMDKNLECKRILTGHTGSLVCLQYDE--RVIIITGSSDSIVRV 331
DB 263 QIHDDVAVTGSDMTLKWICIDKGEVMTYVGHGTGVTWTSOISCGRYIVSGSTDRIVKV 322
QY 332 MDVNTGEMLNLHHCFAVHLRFNNGMMYTCSDRSIAVDMASPTDITLRVYVGHRA 391
DB 333 WSYVDGSLHTLQGHSTVRCMAAGSLIVGSDTILRWVDVSGRHLA---TLHGHA 379
QY 392 AVNVVDEDDKIYVASGDRITIKVNTSTCEFRVTLNGHKGRIACLOQRDR--LVVSGSSD 449
DB 380 AVRQVDFGTIVSGGYDFYVKINNAHTGRCIRLTGTNNHNVSLPSESEKSYICSSGLD 439
QY 450 NTIRLMDI---ECGACLRVLEGHELVRCIFDKMKRIVSGAVDKIKVMDLVALDRAP 506
DB 440 TSIKRWMTTREGGECAVALLQGHSTLSTGMLRGNIILVSCNADSHVRWDI-----H 491
QY 507 ATGLCLFTLVHSGRVFRLO--FDEFQIVSSSHDDITILMD 545
DB 492 EGT-CVHMLSGHRSATITSLQWFGRMVATSSDDGTIVKIMD 530

```

```

RESULT 4
US-08-190-802A-32
Sequence 32, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 265

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fadian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0980
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ. ID NO. 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 779 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CD4 / CD20 protein, Fig. 15
US-08-190-802A-32

```

```

Query Match      13.2%; Score 399; DB 1; Length 779;
Best Local Similarity 24.3%; Pred. No. 2,8e-33;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

```

```

QY 56 TVCLASTAMKTCNVAKTKLANGSTSMIVPKQKLSASYEKK-----ELCVKYFEQW 108
DB 194 TPLATTTKTINN-----NNNIADLIESKDISPEVLSIDEIFSAINNPHAYFK-- 244
QY 109 SESDQVEFEVHLISQMCYHGHINSYLPMLQDPTLAPARGDHAENILSYDAKSL 168
DB 245 -----NLFRIVANNDRELSDLGTLKDNLRDLITSLP---ELSLKFTNLOPED 294
QY 169 LCAAEVCKEWR-VTSDGMLMKLI--ERMVRTDSLMRGLAERRGQGYLFKNRPPDGN 225
DB 295 IISLGVSOQWNNKIIKRSISLMKKLISENFV-----SPKGF 331
QY 226 APPNSFYALPKIIDI-----IET--IESNRCGRHSIORHCRSETSKGYCLOYD 276
DB 332 NSLNLKLSQKPKLSOODRLRLSFLNIFLKNWYMPKVPQPTTLRGHMTSVITCLOPE 391
QY 277 DQKIVSGLRDNTIKIMDKNLECKRILTGHTGSLVCLQYDE--RVIIITGSSDSIVRV 335
DB 392 DNTVITGADKMTIRVDSIKKFLDLQSGHGVYMLKTAHGSILVSGSTDRIVRWMDIK 451
QY 336 TGEMLNTLHHCFAVHLRFNNGMMYTCSDRSIAVDMASPTDITLRVYVGHRAAVV 395
DB 452 KG-----CCT-----HVFEGHNSVRC 468
QY 396 VDEPD---KIYVASGDRITIKVNTSTCEFRVTLNGHKGRIACLOQRDR--LVVSGSSD 428
DB 469 LDIVYKNIKIYVIGSDNLTILHWKLPKRESSVPDHEBHDYPLVFTPEBNPYFVGLWG 528
QY 429 HKRGIACTOYRDLRWVSGSDNTIRLMDIEGACLRVLEGHELVRCIFRDN--KRIVSG 486
DB 529 HMASRVISGHNIVSGSDNTILWVDMQMCIVLTSHTBRIVSTIYDHRKRCISA 588
QY 487 AYDKIRVMDL-----VAALDPRAPAGTL--CLRTLVHSGRVFRLOPEQIVSS 535

```

Db 589 SMDTIRIWDLENINWNGECASYATNSASPCAKILGAMTYLQGHIALVGLLRLSDKFLVSA 648
 QY 536 SHDDTILIMDFLNDPAAQAEPFRSPSRITYY 566
 Db 649 AADGSIRGMD-AND-----YSRKFSY 668

RESULT 5 US-08-477-346-32

; Sequence 32, Application US/08477346
 ; Patent No. 6262023
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron, Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: thereof
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morison & Foerster
 ; STREET: 2000 Pennsylvania Avenue, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/477, 346
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/487, 072
 ; FILING DATE: 07-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 2550-0025.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 779 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
 ; US-08-477-346-32

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best local Similarity 24.3%; Pred. No. 2.8e-33;

Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCIASHTAMKTENCVAKTKLANGTSSMIVPKOKLSASYEKER-----ELCVYFEQW 108
 Db 194 TPLAKTKTINN-----NNNIADLISKDSIISPELSEISATINNINPHAYFK-- 244
 QY 109 SESDVEFEVHILISOMCHQHGHSYIAKPMQORDITLAPRGIDHIAENILSYLDAS 168
 Db 245 -----NLFLRIYVANDRSELSDGLTLDNKLRLDILSLP-----ELSLKTFNYLQFED 294
 QY 169 LCAAEIVCKEWR-VTSDMLWKIL--ERAVRTDSLWGLAERGWQGYLFKNKPPDGN 225
 Db 295 IINSLVSONNMKIIKISLWKKLILISNEY-----SPKGF 331

QY 226 APPNSFYRLYPKIID-----IET--IESNRCGRHSIORHCRSETSKGVYCLQYD 276
 Db 332 NSLNLKLSQRYKPLSOODRLRLSLFLENIFLTKNMWYPRKVPQRTLLRGHMTSVICLOE 391
 QY 277 DOKIVSGLRNTIKIWDKNTLBECKRLTGTGTSVGLQYD-RVIYTGSSDSIVRWYDVN 335
 Db 392 DNYVTGADDMKIRYDSINKKFLQLSGHDGVALKTAHGIIIVSGSTDRIVWIDIK 451
 QY 336 TGEMLNTLLHHCRAVLHLEFRNNGMWTCSKRSIAVWMASTPDLTLRLVGLGHAAYV 395
 Db 452 KG-----CCT-----HVEFGHSTYRC 468
 QY 396 VDPDD-----KIYVASGDKTIKWMNT-----STCE---FVPTLNG 428
 Db 469 LDIVEYKNIKIYVTSGRDNTLHWKLPKSSVPDGHGEHDYPLVFTPEENPYFVGLRG 528
 QY 429 HKRGIAQLQRLVYSGSSDNTIRLMDIEGACALVLEGHELRRCIRFEDN-KRIYSG 486
 Db 529 HMASVTVSGHGNIVYSGSDNTLWYDVAOKCLYILSGHTDRITYSTIYDERRKCTISA 588
 QY 487 AIDGKTKWDL-----VALDPRAPAGTL--CLRTLVHSGRVRLQFDEFOIVSS 535
 Db 589 SMDTIRIWDLENINWNGECASYATNSASPCAKILGAMTYLQGHIALVGLLRLSDKFLVSA 648
 QY 536 SHDDTILIMDFLNDPAAQAEPFRSPSRITYY 566
 Db 649 AADGSIRGMD-AND-----YSRKFSY 668

RESULT 6

; Sequence 3, Application US/08283917
 ; Patent No. 5849557
 ; GENERAL INFORMATION:
 ; APPLICANT: ADACHI, HIDEKI
 ; APPLICANT: TSUJIMOTO, MASAFUMI
 ; APPLICANT: INOUE, KEIZO
 ; APPLICANT: ARAI, HIROYUKI
 ; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 ; TITLE OF INVENTION: AND GENE THEREOF
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
 ; ADDRESS: NEWSTADT, P.C.
 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/283,917
 ; FILING DATE: 03-AUG-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 209943/1993
 ; FILING DATE: 03-AUG-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oblon, No. 5849557man F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 2292-030-0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 409 amino acids
 ; TYPE: amino acid

```

; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
US-08-283-917-3

```

| | | | | |
|-----------------------|--------|--------------------|-------|-----------------------------------|
| Query Match | 11.7%; | Score 354; | DB 2; | Length 409; |
| Best Local Similarity | 29.2%; | Pred. No. 5.8e-29; | | |
| Matches | 87; | Conservative | 57; | Mismatches 110; Indels 44; Gaps 8 |

```

QY 280 IYSGLENDITKIMDNTECKRIILGNHGSVLCODE--RVITGSSDSPTXWMDVATG 337
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 122 MWSASEBATIKWDETGTDEFEKTLKGTIDSVODISPHSKILLASCSAMTKIMLDPGFG 180
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 338 EMINTLIHHCSEAVLHFR--NNGMKVYTCSDKRSIAVMDMASPTDITLRVYLVGHPAAVNV 392
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 ECITMGHNDHNHNVSAIIPNDHDIHVSASRDKTIKMEVDTGYCV---KTFGHEWVRM 236
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 396 V--DEDDKYIVASGORTIKWNTSTCEAFVTLGNHRCIACIÖYDR----- 444
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 239 VRPNODSTLIASCSNDQYIRVAVVATKCKRKLREHNHYECISMAPPESSYSSEATGS 296
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 442 -----LVYSGSSDNTIRLMDIECGACRLVBSGHEELVBCIRFDN--KRIYSGAYD 488
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 299 ETKSKSGPGFELLGSRDRTIKMMDVSTGCMILVQHDMMVWGVLFHSGKFFILSCADD 356
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 490 GKIKWMDVLAALDPRAPACTLRLFLVHSGRVRFLQDEEF--QIYSSHDITILMD 545
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 359 KTLWMDVY-----KKRCKMKTLANHEHFPVSLDFPKHTAIVYVTVGSVDQTVKWE 407
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 7
US-08-961-716-3
: Sequence 3, Application US/08961716
: Patent No. 5880272
: GENERAL INFORMATION:
: APPLICANT: ADACHI, HIDEKI
: APPLICANT: TSUJIMOTO, MASAFUMI
: APPLICANT: INOUE, KEIJO
: APPLICANT: ARAI, HIROYUKI
: TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
: TITLE OF INVENTION: AND GENE THEREOF
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBIJON, SEIVAK, MCCLELLAND, MAIER &
: ADDRESSEE: NEUSTADT, P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961, 716
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/283,917
: FILING DATE: 03-AUG-1994
: APPLICATION NUMBER: JP 209943/1993
: FILING DATE: 03-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Obijon, No. 5880272man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 2292-030-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000

```

1 TELEFAX: (703) 413-2220
2
3 TELE: 248855 OPAT UR
4
5 INFORMATION FOR SEQ. ID NO.:
6
7     SEQUENCE CHARACTERISTICS:
8
9     LENGTH: 409 amino acids
10
11     TYPE: amino acid
12
13     STRANDEDNESS: unknown
14
15     TOPOLOGY: unknown
16
17     MOLECULE TYPE: peptide
18
19     ORIGINAL SOURCE:
20
21     ORGANISM: Bos taurus
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
10
```

| | | | | |
|-----------------------|--------|--------------------|-------|-----------------|
| Query Match | 11.78; | Score 354; | DB 2; | Length 409; |
| Best Local Similarity | 29.28; | Pred. No. 5.8e-29; | | |
| Matches | 87; | Conservative | 57; | Mismatches 110; |
| | | | | Indels 44; |
| | | | | Gaps 8. |

[illegible]

RESULT 8
 US-08-283-917-9
 : Sequence 9, Application US/08283917
 : Patent No. 5849557
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: ADACHI, HIDEKI
 : APPLICANT: TSUTSUMOTO, MASAFUMI
 : APPLICANT: INOUE, KEIZO
 : APPLICANT: ARAI, HIROUKI
 : TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 : TITLE OF INVENTION: AND GENE THEREOF
 : NUMBER OF SEQUENCES: 31
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER &
 : ADDRESSEE: NEUSTADT, P.C.
 : STREET: 1755 S. Jefferson Davis Highway, Suite 400
 : CITY: Arlington
 : STATE: Virginia
 : COUNTRY: U.S.A.
 :
 : ZIP: 22202
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentln Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/283,917
 : FILING DATE: 03-AUG-1994
 : CLASSIFICATION: 435
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 209943/1993
 : FILING DATE: 03-AUG-1993
 :
 : ATTORNEY/AGENT INFORMATION:
 :
 : NAME: OBION, NO. 5849557/man F.

```

1      RESULT 10
2      US-08-190-802A-51
3      ; Sequence 51, Application US/08190802A
4      ; Patent No. 5519003
5      ;
6      ; GENERAL INFORMATION:
7      ;
8      ; APPLICANT: Moench-Rosen, Darla
9      ;
10     ; APPLICANT: Ron, Dorit
11     ;
12     ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
13     ;
14     ; TITLE OF INVENTION: Thereof
15     ;
16     ; NUMBER OF SEQUENCES: 265
17     ;
18     ; CORRESPONDENCE ADDRESS:
19     ;
20     ; ADDRESSEE: Delhinger & Associates
21     ;
22     ; STREET: P.O. Box 60850
23     ;
24     ; CITY: Palo Alto
25     ;
26     ; STATE: CA
27     ;
28     ; COUNTRY: USA
29     ;
30     ; ZIP: 94306-0850
31     ;
32     ; COMPUTER READABLE FORM:
33     ;
34     ; MEDIUM TYPE: Floppy disk
35     ;
36     ; COMPUTER: IBM PC compatible
37     ;
38     ; OPERATING SYSTEM: PC-DOS/MS-DOS
39     ;
40     ; SOFTWARE: Patent In Release #1.0, Version #1.25
41     ;
42     ; CURRENT APPLICATION DATA:
43     ;
44     ; APPLICATION NUMBER: US/08/190,802A
45     ;
46     ; FILING DATE: 01-FEB-1994
47     ;
48     ; CLASSIFICATION: 530
49     ;
50     ; ATTORNEY/AGENT INFORMATION:
51     ;
52     ; NAME: Fabian, Gary R.
53     ;
54     ; REGISTRATION NUMBER: 33,875
55     ;
56     ; REFERENCE/DOCKET NUMBER: 8600-0139

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: L1S1 (human), Fig. 34
US-08-190-802A-51

Query Match      11.2%: Score 339.5; DB 1; Length 409;
Best Local Similarity 28.9%: Pred. No. 2e-27;
Matches 86; Conservative 57; Mismatches 110; Indels 45; Gaps 9;

QY 280 IVSGLDNTIKIKDKNTLECKRIITGHTGTVLCLOYDE--RVITGSSDSSTVRVMDVNTG 337
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 MVSASEDATIKWVDYETGDFERTLKHTDSYQDISFDHSGKLASCSADMTIKIMDFOGF 182
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 338 EMLNTLIHCEAVLHLRF--NNGMVTGSKDRSIAVMDASPTDITLRVLYGHRRAVNV 395
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 183 ECIRTMGHNDHNVSSVAIMPNGDHIYASASRDKITKMEVOTGYCV---KFTFGHREVM 239
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 396 V--DFDDKIYVSGDRITIKVNTSTCEVFTLNHGKRGIAQLQYDR----- 441
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 240 VRPDQGTILASCSNDQTVRWVAVATKECKALREHEHVEICISWAPSSYSISSEATGS 299
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 442 -----LVVSSSDNTIRLMDIEGACLRVLEGEHELVRICIFDN--KRIYGAID 489
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 ETKKSGKPGFLLSGSDKT-KMMDVSTGMCIMTLVGHNDVWVGLVLFHSGGKFTILSCAD 358
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 490 GKIKWMDLVALDPPRAGTLCRLTVEHSGVRFLODFEF--QIVSSSHDTILIMD 545
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 359 KTLRWMDY-----KNKRCMKTLNAHEHFVTSIDFHTAPYVVTGSDQTVKWE 407
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-08-477-346-51
; Sequence 51, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,346
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,072
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959

```

```

REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: L1S1 (human), Fig. 34
US-08-477-346-51

Query Match      11.2%: Score 339.5; DB 4; Length 409;
Best Local Similarity 28.9%: Pred. No. 2e-27;
Matches 86; Conservative 57; Mismatches 110; Indels 45; Gaps 9;

QY 280 IVSGLDNTIKIKDKNTLECKRIITGHTGTVLCLOYDE--RVITGSSDSSTVRVMDVNTG 337
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 MVSASEDATIKWVDYETGDFERTLKHTDSYQDISFDHSGKLASCSADMTIKIMDFOGF 182
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 338 EMLNTLIHCEAVLHLRF--NNGMVTGSKDRSIAVMDASPTDITLRVLYGHRRAVNV 395
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 183 ECIRTMGHNDHNVSSVAIMPNGDHIYASASRDKITKMEVOTGYCV---KFTFGHREVM 239
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 396 V--DFDDKIYVSGDRITIKVNTSTCEVFTLNHGKRGIAQLQYDR----- 441
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 240 VRPDQGTILASCSNDQTVRWVAVATKECKALREHEHVEICISWAPSSYSISSEATGS 299
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 442 -----LVVSSSDNTIRLMDIEGACLRVLEGEHELVRICIFDN--KRIYGAID 489
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 ETKKSGKPGFLLSGSDKT-KMMDVSTGMCIMTLVGHNDVWVGLVLFHSGGKFTILSCAD 358
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 490 GKIKWMDLVALDPPRAGTLCRLTVEHSGVRFLODFEF--QIVSSSHDTILIMD 545
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 359 KTLRWMDY-----KNKRCMKTLNAHEHFVTSIDFHTAPYVVTGSDQTVKWE 407
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-08-190-802A-66
; Sequence 66, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P. O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fadian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:

```

RESULT 14
US-08-190-802A-52
Sequence 52, Application US/08190802A
Patent No. 5518003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dellinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 52:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 422 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: MD6, Fig. 35
 US-08-190-802A-52

Query Match 10.5%; Score 318; DB 1; Length 422;
 Best Local Similarity 24.4%; Pred. No. 4.1e-25;
 Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

QY 103 KYFEQMSSEDVDF-----VEHLISQMCHYGHINSYLMKMLORDPITALPAR 151
 DB 4 KDFETWLDNISVTFSLMDLQKNETLDHLISLGAVALRHLISNNLETLKRDPLKLP- 62
 QY 152 GLDHAENILSYLDAKSLCAELVCKEYRVSQGLMKLIERVATDSLMRGLAERG 211
 DB 63 ---ELSFYLLKWLDPQTLTCLVSKQKRVIS-----ACTEVMQACKNLG 106
 QY 212 WQOYLKKNPPGNAAPPNSFYRALYPKIIODIETIESNMKGRHSIORHCRSETSKGY 271
 DB 107 W-----QIDDSVQDSILHMKRYLKAIRKOLED-----HEAFETSS--- 143
 QY 272 CLQYDDQKIVSGLRDNTIKIMDKNTLECKRLTGHGTVSLQYDERVITGSSDSTVR 331
 DB 144 -----LIGHSARVYALYKDGCLCTGSDLSAKL 172
 QY 332 WDVNTGEMLNTL-IHCEAVLHLRFNNGMMVTCSDRSIAVMDASPTDITLRVLVGR 390
 DB 173 WDVSTGQCVYGIQHTCAAV---KFDEQKLVTSFQNTVAQWEMSSGARTQHR---GHT 226
 QY 391 AAVNVDPFDK--YIVSASGDRITKVMNTSTCEPVRTLNGHKRG-----ACIQ 437
 DB 227 GAVFSVYSDLDLIVSGSADFAVKWALSAGTCLNLTGHTTEWTKVYLQCKVKSLH 286
 QY 438 YRDLVYSSGSDNTIRLW---DIEGACLRVLEGHEELVRCIR---FNNKRIYS--- 485
 DB 287 SPQDYILSADKYEIKIPIRGREINC-KCLKTLYSVSDRSICLOPRLHFDGKIYVCSAL 345
 QY 486 GAYDGKIKVMDLVAALDPAPAGTLCRLTVLHSGRVRFLOFD 528
 DB 346 GLYQWDFASYDILRVIKTPEVANLALL-----GFGDVFALLFD 383

RESULT 15
 US-08-477-346-52
 Sequence 52, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 52:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 422 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: MD6, Fig. 35
 US-08-477-346-52

Query Match 10.5%; Score 318; DB 4; Length 422;
 Best Local Similarity 24.4%; Pred. No. 4.1e-25;
 Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

QY 103 KYFEQMSSEDVDF-----VEHLISQMCHYGHINSYLMKMLORDPITALPAR 151
 DB 4 KDFETWLDNISVTFSLMDLQKNETLDHLISLGAVALRHLISNNLETLKRDPLKLP- 62
 QY 152 GLDHAENILSYLDAKSLCAELVCKEYRVSQGLMKLIERVATDSLMRGLAERG 211
 DB 63 ---ELSFYLLKWLDPQTLTCLVSKQKRVIS-----ACTEVMQACKNLG 106
 QY 212 WQOYLKKNPPGNAAPPNSFYRALYPKIIODIETIESNMKGRHSIORHCRSETSKGY 271
 DB 107 W-----QIDDSVQDSILHMKRYLKAIRKOLED-----HEAFETSS--- 143
 QY 272 CLQYDDQKIVSGLRDNTIKIMDKNTLECKRLTGHGTVSLQYDERVITGSSDSTVR 331
 DB 144 -----LIGHSARVYALYKDGCLCTGSDLSAKL 172
 QY 332 WDVNTGEMLNTL-IHCEAVLHLRFNNGMMVTCSDRSIAVMDASPTDITLRVLVGR 390
 DB 173 WDVSTGQCVYGIQHTCAAV---KFDEQKLVTSFQNTVAQWEMSSGARTQHR---GHT 226
 QY 391 AAVNVDPFDK--YIVSASGDRITKVMNTSTCEPVRTLNGHKRG-----ACIQ 437
 DB 227 GAVFSVYSDLDLIVSGSADFAVKWALSAGTCLNLTGHTTEWTKVYLQCKVKSLH 286
 QY 438 YRDLVYSSGSDNTIRLW---DIEGACLRVLEGHEELVRCIR---FNNKRIYS--- 485
 DB 287 SPQDYILSADKYEIKIPIRGREINC-KCLKTLYSVSDRSICLOPRLHFDGKIYVCSAL 345
 QY 486 GAYDGKIKVMDLVAALDPAPAGTLCRLTVLHSGRVRFLOFD 528
 DB 346 GLYQWDFASYDILRVIKTPEVANLALL-----GFGDVFALLFD 383

Search completed: May 8, 2002, 10:50:44
 Job time: 111 sec

